

Programa de Doctorat en Ciències Experimentals i Tecnologies

Línia d'investigació d'Enginyeria i Tecnologia

Doctorat Industrial:

Institut Català de la Salut i Universitat de Vic – Universitat Central de Catalunya

**AVALUACIÓ DE L'IMPACTE DE LA INTEL·LIGÈNCIA ARTIFICIAL
EN L'ÀMBIT DE LA SALUT I VALIDACIÓ D'UN ALGORISME
PER LA LECTURA DE RADIOGRAFIES DE TÒRAX EN L'ÀMBIT
D'ATENCIÓ PRIMÀRIA**

Tesi presentada per

Queralt Miró Catalina

Departament d'Enginyeries

Facultat de Ciències, Tecnologia i Enginyeries

Universitat de Vic – Universitat Central de Catalunya

Directors:

Dr. Josep Vidal Alaball

Dr. Jordi Solé Casals

Vic, Barcelona, 2024

El 80% de l'èxit es basa, senzillament, en insistir.

A.

AGRAÏMENTS

Fent un repàs de tot el que m'ha ajudat a arribar fins aquí, recordo el primer cop que em vaig començar a plantejar això de fer un doctorat. Va ser al acabar la carrera i començar a treballar a la Universitat Internacional de Catalunya:

- Com i què deu ser això de fer un doctorat ? - Vaig pensar.

Un nou món es va obrir davant meu. Allà vaig descobrir què era fer recerca, estar a una unitat d'investigació i docència, i compartir el meu dia a dia amb doctors/es i futurs doctors/es. Als pocs mesos d'arribar, amb ganes (i perquè no, certs dubtes) em vaig matricular a un màster per compaginar amb la feina. *Oficial sobretot*, em deien, i ho argumentaven amb un *per si de cas, per no tancar-te portes*.

Pocs mesos després d'acabar el màster i amb una mica més d'experiència i formació, vaig iniciar un nou capítol professional. Vaig començar a treballar a on soc ara, a la Unitat de Suport a la Recerca de l'ICS. Aquí, vaig tenir la sort de trobar-me amb el Josep, el cap de la Unitat i un dels directors de la tesi. Sempre amb un sí assegurat, disposat a fer més, a avançar, a ajudar-nos a ser millors. Gràcies a les seves ganes, ambició i inquietuds per la recerca vaig iniciar aquest projecte de la tesi ara fa poc més de dos anys.

Al ser un doctorat industrial, necessitàvem un altre director de tesi d'una universitat i gràcies al vincle del Josep amb la Universitat de Vic – Universitat Central de Catalunya, vam contactar amb el Jordi. Quina sort vaig tenir, no podríem haver triat millor. Tots dos sempre disponibles, sumant i sense posar través al camí.

Ara, a punt de tancar un nou capítol, puc dir que durant el camí he après a fer sempre les coses amb ganes, siguin quines siguin les circumstàncies, per poder gaudir del camí celebrant els passos positius i aprenent dels negatius. Per això, mirant enrere, sento enyorança i emoció de tot el viscut i estic preparada per tancar una etapa, expectant per obrir-ne una altre i sobretot agraïda d'haver arribat fins aquí.

Només puc donar les gràcies a tots i cada un de vosaltres per haver format part del camí. A Oxipit per deixar-nos provar el seu algorisme, als companys i amics meravellosos de la Universitat Internacional de Catalunya per ajudar-me a fer el primer pas. A la Unitat de Suport a la Recerca de l'ICS per estar envoltada d'un equip brillant de professionals i sobretot, de grans companys i companyes, alguns d'ells indispensables per arribar fins aquí, fent el camí més fàcil, més divertit i amb la capacitat sempre d'ajudar-me a veure més enllà.

I en especial a la meva família, per ser molt més que família. A la meva mare, per tenir sempre una branca aquí o allà on poder fer parada, agafar aire i seguir. Al meu pare, per guiar-me i cuidar-me sempre des de les estrelles. I als dos ulls marrons i brillants que m'acompanyen des de fa temps, a les bones i a les no tant bones, i em miren tan bonics dient-me: endavant, junts podem amb tot.

Moltes gràcies!

RESUM

El primer cop que es va parlar del terme d'intel·ligència artificial (IA) va ser al 1956. Des d'aleshores, i gràcies a l'actual volum de dades generades, a la capacitat d'emmagatzemar-les i el programari per treballar-les, el seu impacte no ha parat d'evolucionar i créixer fins estar present en gairebé tots els aspectes del dia a dia de la població. Un dels àmbits on hi ha major esperança en l'aplicació de la IA és en l'àmbit de la salut. Es considera que pot impactar directament en la millora de la salut de la població ajudant a la fiabilitat i certesa en la presa de decisions, millorant la confiança en el diagnòstic, l'eficiència en el flux de treball i permeten una atenció més personalitzada, entre d'altres. Tant és així que, en els darrers anys, al nombre de publicacions indexades a revistes científiques sobre IA en salut ha passat del 6% entre el 1995 i el 2010 al 43% entre el 2014 i el 2019.

Concretament, una de les especialitats on hi ha un major nombre d'aplicacions d'IA actuals i un ventall d'aplicacions futures més gran és en la radiologia, una especialitat que es troba present en el 50% dels pacients atesos a urgències i en el 70% dels pacients hospitalitzats i que, malauradament, en els darrers anys el nombre d'especialistes està disminuint i s'està evidenciant una important manca de professionals.

Tot i que actualment hi ha diverses aplicacions d'IA per la lectura de tòrax, la majoria d'elles no tenen una validació externa del seu funcionament per poder assegurar-ne el seu funcionament al món real. A més, cal posar èmfasi en les percepcions i coneixements tan de la població general com dels professionals sanitaris sobre l'ús d'eines d'IA en l'àmbit de la salut. De fet, alguns estudis ja han observat sentiments positius i esperançadors amb l'ajuda d'eines d'IA en l'àmbit de la salut, però també preocupacions i desconeixements que cal conèixer i solucionar per poder assegurar una bona acceptació i ús d'aquestes.

Per tot això, els objectius de la present tesi són: 1) conèixer la percepció i coneixement dels professionals sanitaris d'atenció primària sobre l'ús de la IA i el seu impacte en la radiologia, 2) conèixer la percepció i coneixement de la població sobre l'ús de la IA en la radiologia i 3) realitzar una validació externa amb dades prospectives en l'àmbit d'atenció primària d'un algorisme d'IA per la lectura de radiografies de tòrax. Per donar resposta als objectius plantejats la present tesi forma un compendi de 4 articles publicats a revistes científiques.

Per tal de donar resposta als dos primers objectius s'ha realitzat dos estudis transversals a través de dues enquestes validades, una pels professionals sanitaris i una altre per la població general. L'enquesta dels professionals l'han pogut respondre tots els professionals d'infermeria i medicina de l'Institut Català de la Salut de la Catalunya Central i l'enquesta de la població general l'han pogut respondre tots els ciutadans de Catalunya majors de 18 anys. Per tal de donar resposta al tercer objectiu de la validació externa de l'algorisme s'ha realitzat

un estudi prospectiu comparant el diagnòstic proposat per l'algorisme d'IA i el diagnòstic del radiòleg de referència, considerant aquest últim com el *gold standard*.

En relació al primer objectiu, l'enquesta s'ha respost per 301 professionals sanitaris d'atenció primària. La majoria ha indicat conèixer el concepte d'IA tot i mostrar dubtes sobre si l'utilitzen o no a la consulta, i han indicat una manca de formació en IA i un gran interès en rebre'n. S'ha observat una percepció positiva de l'impacte de la IA en salut, especialment en la millora de les decisions clíniques i en la millora de la salut de la població, però existeix certa preocupació per no sentir-se correctament preparats per la implementació de la IA en les consultes. Aquestes percepcions han sigut superiors en els professionals amb coneixements o interès sobre la IA.

En relació al segon objectiu, s'ha obtingut 379 respostes i la majoria ha indicat que coneixia el concepte d'IA. Tot i això, s'ha observat un sentiment lleugerament negatiu respecte l'aplicació de la IA en salut i en la radiologia. L'àmbit on s'ha mostrat més preocupació ha sigut en la possible pèrdua d'interacció social, en la necessitat d'estar informat, la desconfiança i la falta de responsabilitat de la IA. A més, en les persones amb nivells d'estudis baixos, sense coneixements sobre IA, d'edats majors a 65 o de sexe femení s'ha observat percepcions més negatives respecte l'aplicació de la IA en relació a la resta. En canvi, s'ha observat certa esperança en l'eficiència que pot oferir la IA.

En relació al tercer objectiu, s'ha realitzat una validació externa amb 278 imatges i informes. El 51.8% d'aquestes imatges eren sense alteracions radiològiques segons l'informe del radiòleg. L'algorisme ha obtingut un valor de l'*accuracy* de 0.95 (IC 95% 0.92; 0.98), de la sensibilitat de 0.48 (IC 95% 0.30; 0.66) i de l'especificitat de 0.98 (IC 95% 0.97; 0.99). Les patologies on l'algorisme ha tingut més sensibilitat han sigut els implants externs, d'abdomen superior i cardíaca i/o valvular. En canvi, les patologies on l'algorisme ha tingut menys sensibilitat han sigut mediastí, vasos i òssia. Amb això, l'algorisme ha demostrat ser útil per poder identificar radiografies amb o sense alteracions però s'ha de seguir entrenant per poder ser una bona eina de suport i ajuda en el diagnòstic radiològic, a nivell general i en determinades patologies en concret.

RESUMEN

La primera vez que se habló del término de inteligencia artificial (IA) fue en 1956. Desde entonces, y gracias al actual volumen de datos generados, a la capacidad de almacenarlos y al software para trabajarlos, su impacto no ha dejado de evolucionar y crecer hasta estar presente en casi todos los aspectos del día a día de la población. Uno de los ámbitos donde hay mayor esperanza en la aplicación de la IA es en el ámbito de la salud. Se considera que puede impactar directamente en la mejora de la salud de la población ayudando a la fiabilidad y certeza en la toma de decisiones, mejorando la confianza en el diagnóstico, la eficiencia en el flujo de trabajo y permitiendo una atención más personalizada, entre otros. Tanto es así que, en los últimos años, el número de publicaciones indexadas en revistas científicas sobre IA en salud ha pasado del 6% entre 1995 y 2010 al 43% entre 2014 y 2019.

Particularmente, una de las especialidades donde hay un mayor número de aplicaciones de IA actuales y un abanico de aplicaciones futuras más grande es en la radiología, una especialidad que se encuentra presente en el 50% de los pacientes atendidos en urgencias y en el 70% de los pacientes hospitalizados y que desafortunadamente, en los últimos años el número de especialistas está disminuyendo y se está evidenciando una importante falta de profesionales. Por todo esto, actualmente hay varias aplicaciones de IA para la lectura de tórax, pero la mayoría de ellas no tienen una validación externa de su funcionamiento para poder dar validez a sus capacidades y asegurar su funcionamiento en el mundo real.

Además, es necesario poner énfasis en las percepciones y conocimientos tanto de la población general como de los profesionales sanitarios sobre el uso de herramientas de IA en el ámbito de la salud. De hecho, algunos estudios ya han observado sentimientos positivos y esperanzadores con la ayuda de herramientas de IA en el ámbito de la salud pero también preocupaciones y desconocimientos que son necesarios conocer y solucionar para poder asegurar una buena aceptación y uso de estas.

Por todo esto, los objetivos de la presente tesis son: 1) conocer la percepción y conocimiento de los profesionales sanitarios de atención primaria sobre el uso de la IA y su impacto en la radiología, 2) conocer la percepción y conocimiento de la población sobre el uso de la IA en la radiología y 3) realizar una validación externa con datos prospectivos en el ámbito de

atención primaria de un algoritmo de IA para la lectura de radiografías de tórax. Para dar respuesta a los objetivos planteados la presente tesis forma un compendio de 4 artículos publicados en revistas científicas.

Para dar respuesta a los dos primeros objetivos, se han realizado dos estudios transversales a través de dos encuestas validadas, una para los profesionales sanitarios y otra para la población general. La encuesta de los profesionales la han podido responder todos los profesionales de enfermería y medicina del Institut Català de la Salut de la Catalunya Central y la encuesta de la población general la han podido responder todos los ciudadanos de Catalunya mayores de 18 años. Para dar respuesta al tercer objetivo de la validación externa del algoritmo se ha realizado un estudio prospectivo comparando el diagnóstico propuesto por el algoritmo de IA y el diagnóstico del radiólogo de referencia, considerando éste último como el gold standard.

En relación al primer objetivo, la encuesta ha sido respondida por 301 profesionales sanitarios de atención primaria. La mayoría ha indicado conocer el concepto de IA aunque muestran dudas sobre si la utilizan o no en la consulta, y han indicado una falta de formación en IA y un gran interés en recibirla. Se ha observado una percepción positiva del impacto de la IA en salud, especialmente en la mejora de las decisiones clínicas y en la mejora de la salud de la población, pero existe cierta preocupación por no sentirse correctamente preparados para la implementación de la IA en las consultas. Además, estas percepciones han sido superiores en los profesionales con conocimientos o interés sobre la IA.

En relación al segundo objetivo, se han obtenido 379 respuestas y la mayoría ha indicado que conocía el concepto de IA. Sin embargo, se ha observado un sentimiento ligeramente negativo respecto a la aplicación de la IA en salud y en la radiología. El ámbito donde se ha mostrado más preocupación ha sido en la posible pérdida de interacción social, en la necesidad de estar informado, la desconfianza y la falta de responsabilidad de la IA. En cambio, se ha observado cierta esperanza en la eficiencia que puede ofrecer la IA. Además, en las personas con niveles de estudios bajos, sin conocimientos sobre IA, de edades mayores a 65 años o de sexo femenino se ha observado percepciones más negativas respecto a la aplicación de la IA en relación al resto.

En relación al tercer objetivo, se ha realizado una validación externa con 278 imágenes e informes siendo el 51.8% de estas imágenes sin alteraciones radiológicas según el informe del radiólogo. El algoritmo ha obtenido un valor de la precisión de 0.95 (IC 95% 0.92; 0.98), de la sensibilidad de 0.48 (IC 95% 0.30; 0.66) y de la especificidad de 0.98 (IC 95% 0.97; 0.99). Las patologías donde el algoritmo ha tenido más sensibilidad han sido los implantes externos, de abdomen superior y cardíaca y/o valvular. En cambio, las patologías donde el algoritmo ha tenido menos sensibilidad han sido mediastino, vasos y ósea. Con esto, el algoritmo ha demostrado ser útil para poder identificar radiografías con o sin alteraciones pero se debe seguir entrenando para poder ser una buena herramienta de apoyo y ayuda en el diagnóstico radiológico, a nivel general y en determinadas patologías en concreto.

ABSTRACT

The term 'artificial intelligence' (AI) was first discussed in 1956. Since then, there has been a remarkable increase in the volume of data generated, and the ability to store and work on it has evolved. Consequently, IA has become a part of nearly every aspect of the population's daily lives.

Health is a promising area for AI application. It has the potential to lead to a significant improvement in the health of the population, by facilitating reliability and certainty in decision-making, improving confidence in the diagnosis, reducing workflow efficiency, and allowing for more personalized care, among other things. As a result, the number of articles indexed in scientific journals concerning AI in health has grown from 6% (1995-2010) to 43% (2014-2019) in just a few years.

In particular, radiology is a specialty with a greater number of current AI applications and a greater range of future applications. This specialty that is present in 50% of patients treated in emergency rooms and in 70% of hospitalized patients. Unfortunately, in recent years, the number of specialists has decreased and a significant lack of professionals is becoming evident.

Although there are multiple AI applications for chest reading, most of them lack external validation to guarantee their real-world performance. In addition, it is necessary to emphasize the perceptions and knowledge of both the general population and health professionals regarding the use of AI tools in the healthcare field. In fact, some studies have already reported positive and hopeful feedback using AI tools in this field. Despite this, there are still concerns and lack of awareness that need to be addressed to ensure a good acceptance and utilization of these.

In light of this, the objectives of this thesis are: 1) to understand the perception and knowledge of primary care health professionals regarding the use of AI and its impact on radiology, 2) to acquire knowledge about how the public perceives and understands the use of AI in radiology and 3) to perform an external validation with prospective data in the primary care setting of an AI algorithm for reading chest radiographs. To respond to the stated objectives, this thesis is a compilation of four articles published in scientific journals.

In order to respond to the first two objectives, two cross-sectional studies were carried out using two validated surveys, one for health professionals and the other for the general population. The survey for professionals was answered by all the nursing and medical professionals of the Catalan Institute of Health of Central Catalonia, and the survey for the general population was answered by all the citizens of Catalonia over 18 years of age. To achieve the third objective of the external validation of the algorithm, a prospective study

was carried out comparing the diagnosis proposed by the AI algorithm and the diagnosis of the reference radiologist, the latter being considered the gold standard.

In relation to the first objective, the survey was answered by 301 primary care health professionals. Although the majority reported knowing AI, they expressed doubts about whether or not they would use it in the consultation, and expressed a lack of AI training and a strong desire to receive it. AI has been observed to have a positive impact on health, particularly in the improvement of clinical decisions and the health of the population. There is some uncertainty about not feeling adequately prepared for the implementation of AI in consultations. Professionals with knowledge or interest in AI showed a greater tendency to have these perceptions.

To answer the second objective, 379 responses were obtained and the majority of the population indicated that they were familiar with the concept of artificial intelligence. However, there was a slight negative perception about the use of AI in health and radiology. The concerns were related to the potential loss of social interaction, the need to be informed, distrust, and the lack of responsibility of AI. In addition, these perceptions were more common among individuals with low levels of education, no previous knowledge about AI, over the age of 65, or women. Conversely, the efficiency that AI can offer was perceived as a source of hope.

An external validation was conducted for the third objective, involving 278 images and reports. According to the radiologist's report, 51.8% of these images were not altered in any way. The algorithm obtained an accuracy value of 0.95 (CI 95% 0.92, 0.98), a sensitivity of 0.48 (CI 95% 0.30, 0.66) and a specificity of 0.98 (CI 95% 0.97, 0.99). The algorithm was able to detect external, upper abdominal, cardiac, and/or valvular implants at the highest level of sensitivity. The algorithm's sensitivity was decreased for mediastinum, vessels, and bone pathologies. The algorithm has proven to be useful in identifying radiographs with or without alterations. It requires further training to be an effective tool for supporting and aiding in radiological diagnosis, both at a general level and for certain specific pathologies.

Abreviatures

ACMO: Agrupació de Ciències Mèdiques a Osona

AIMA: Artificial Intelligence, a Modern Approach

ANN: Artificial Neural Networks

AP: Anteroposterior

BMJ: British Medical Journal

CAD: Computer Assisted Diagnosis

CAP: Centre d'Atenció Primària

CASNET: Causal-Associationnal Network

CE: Comunitat Europea

CNN: Convolutional Neural Network

CoMB: Col·legi de Metges de Barcelona

DICOM: Digital Imaging and Communication in Medicine

EE.UU: Estats Units

FDA: Food and Drug Administration

FEC: Facultat d'Empresa i Comunicació

IA: Intel·ligència Artificial

IC: Interval de Confiança

ICS: Institut Català de la Salut

IDIAP: Institut Universitari d'Investigació d'Atenció Primària

IoT: Internet de les Coses

JMIR: Journal of Medical Internet Research

MVS: Milisievert

PA: Posteroanterior

Q1: Primer Quartil

Q2: Segon Quartil

QR: Quick Response

RAE: Real Academia Espanyola

RX: Radiografia

S3VM: Semi-Supervised Support Vector Machine

SD: Standard Deviation

SDPI: Servei de Diagnòstic per la Imatge

SVM: Support Vector Machine

TC: Tomografia Computada

UE: Unió Europea

Índex de figures

Figura 1. Mapa conceptual de les aplicacions i tècniques que utilitza la IA.

Figura 2. Funcionament del perceptró simple.

Figura 3. Funcionament del perceptró multicapa.

Figura 4. Funcionament de processament d'una imatge a través de les CNN.

Figura 5. Exemple d'ús de l'aplicació per a la lectura d'imatge.

Índex

1. TESI PER COMPENDI D'ARTICLES	1
1.1 Publicació d'articles a revistes científiques	1
1.2 Difusió dels resultats de la tesi a congressos o jornades.....	5
1.3 Difusió dels resultats de la tesi als mitjans de comunicació	6
1.4 Ajuts i premis obtinguts.....	8
2. INTRODUCCIÓ.....	10
1.1 Origen i evolució de la intel·ligència artificial	10
1.2 Definició de la intel·ligència artificial	12
1.2.1 Mètodes que utilitza la intel·ligència artificial.....	14
1.2.1.1 Aprenentatge automàtic	14
1.2.1.2 Aprenentatge profund	16
1.2.1.3 Xarxes neuronals convolucionals.....	16
1.3 Intel·ligència artificial en la societat	18
1.3.2 Percepcions i coneixements de la ciutadania.....	19
1.4 Intel·ligència artificial en l'àmbit de la salut	20
1.4.1 Percepcions i coneixements dels usuaris i professionals sanitaris	22
1.5 Radiologia	24
1.5.1 Radiografia de tòrax	26
1.5.2 Aplicació de la intel·ligència artificial a la radiologia	27
1.6 Aplicacions actuals d'intel·ligència artificial per a la lectura de radiografies de tòrax	28
1.6.1 Oxipit	30
1.7 Regulació i legislació de productes i dispositius d'intel·ligència artificial.....	33
1.8 Validació clínica d'algorismes d'intel·ligència artificial.....	35
3. HIPÒTESIS I OBJECTIUS	39
3.1 Hipòtesis	39
3.2 Objectius	39
3.2.1 Objectius generals.....	39
3.2.2. Objectius específics	40
4. MATERIAL I MÈTODES	42
4.1 Enquesta validada a la ciutadania.....	42

4.2 Enquesta validada als professionals	44
4.3 Estudi prospectiu per la validació de l'algorisme.....	45
5. Articles de la tesi doctoral	49
5.1 Article 1: Developing an Artificial Intelligence Model for Reading Chest X-rays: Protocol for a Prospective Validation Study.....	49
5.2 Article 2: Knowledge and perception of primary care healthcare professionals on the use of artificial intelligence as a healthcare tool	56
5.3 Article 3: Knowledge and Perception of the Use of AI and its Implementation in the Field of Radiology: Cross-Sectional Study	68
5.4 Article 4: Real-world testing of an artificial intelligence algorithm for the analysis of chest X-rays in primary care settings	80
6. DISCUSSIÓ.....	94
7. LIMITACIONS	111
8. FUTURES LÍNIES D'INVESTIGACIÓ.....	114
9. CONCLUSIONS	116
10. BIBLIOGRAFIA	119
11. ANNEXES.....	136
Annex I Enquesta original ciutadania.....	136
Annex II Enquesta traduïda al català ciutadania.....	164
Annex III Cartell informatiu sobre l'enquesta a la ciutadania.....	169
Annex IV Enquesta original professionals.....	171
Annex V Enquesta traduïda professionals	175
Annex VI Llistat de patologies detectades amb l'algorisme d'IA i agrupació.....	178

1. TESI PER COMPENDI D'ARTICLES

La present tesi doctoral s'ha realitzat per compendi d'articles, i consta de quatre articles publicats en revistes científiques indexades. S'ha desenvolupat a temps parcial gràcies a una beca del programa de Doctorats Industrials del Departament de Recerca i Universitats de la Generalitat de Catalunya. Aquest programa té per objectiu donar la capacitat de desenvolupar un projecte de recerca entre una empresa o institució en representació de l'entorn empresarial, i una universitat o centre de recerca en representació de l'entorn acadèmic. D'aquesta manera, els doctorands actuen com a ponts de transferència de coneixement entre l'empresa i la universitat o centre de recerca. Així doncs, l'Institut Català de la Salut n'ha sigut l'entorn empresarial i la Universitat de Vic – Universitat Central de Catalunya n'ha sigut l'entorn acadèmic.

1.1 Publicació d'articles a revistes científiques

A continuació s'exposa una breu descripció i informació de la publicació dels quatre articles publicats:

Developing an Artificial Intelligence Model for Reading Chest X-rays: Protocol for a Prospective Validation Study. Miró Catalina, Queralt, Fuster-Casanovas, A., Solé-Casals, J., & Vidal-Alaball, J. (2022). *JMIR Research Protocols*, 11(11), e39536.

La revista *JMIR Research Protocols* té un factor d'impacte al 2023 de 1.7, ocupa la posició Q2 segons *Web of Science* i es troba indexada a *Pubmed* i *Scopus*. Inicialment es centrava en articles de *eHealth*, però actualment publica protocols, propostes d'estudi i resultats preliminars de proves pilot o investigacions formatives de qualsevol àrea de la medicina relacionada amb la salut i la innovació tecnològica. L'article ha estat citat 11 vegades a maig de 2024.

Objectiu: desenvolupament del protocol per realitzar una validació externa amb dades prospectives de l'algorisme d'IA *ChestEye* com a eina de suport al diagnòstic per la lectura de radiografies de tòrax en l'atenció primària. Es vol validar l'algorisme per a una sola patologia i per a múltiples patologies avaluant-ne la precisió, la sensibilitat i l'especificitat.

Resultats esperats: s'espera reclutar la mostra d'estudi durant sis mesos i obtenir proves suficients per demostrar que l'ús de la IA en la lectura de radiografies de tòrax pot ser una bona eina d'ajuda al diagnòstic. El nombre de radiòlegs és cada vegada menor i, per tant, és necessari desenvolupar i validar eines de suport pels professionals que sovint han d'interpretar radiografies de tòrax.

Knowledge and perception of primary care healthcare professionals on the use of artificial intelligence as a healthcare tool. Miró Catalina, Queralt, Fuster-Casanovas, A., Vidal-Alaball, J., Escalé-Besa, A., Marin-Gomez, F. X., Femenia, J., & Solé-Casals, J. (2023). *Digital Health*, 9, 20552076231180511.

La revista *Digital Health* té un factor d'impacte al 2023 de 3.9, ocupa la posició Q1 segons *Journal Impact Factor* (Clarivate, 2023) i es troba indexada a *Pubmed* i *Web of Science*. Publica articles científics sobre la revolució digital en salut amb l'objectiu de crear ponts entre l'evolució de la tecnologia i la informàtica, la medicina, la salut i l'atenció sanitària. L'article ha estat citat 9 vegades a maig de 2024.

Objectiu: descriure la percepció dels professionals d'atenció primària i el seu coneixement sobre l'ús de la IA com a eina de suport i el seu impacte en la radiologia a través d'una enquesta validada, per conèixer les principals fortaleses i preocupacions.

Resultats: l'enquesta es va enviar a 1.068 professionals sanitaris i es van obtenir 301 respostes. El 85.7% va indicar que entenia el concepte de la IA però hi havia discrepàncies en el seu ús, el 65.8% va indicar que no havia rebut formació sobre la IA i el 91.4% que els agradaria rebre'n. La puntuació mitjana per l'impacte positiu de la IA als professionals va ser de 3.62 punts sobre 5 ($sd = 0.72$), amb major puntuació dels qui havien rebut prèviament

formació sobre IA o presentaven interès sobre la IA. La puntuació mitjana de la correcta preparació dels professionals per la aplicació de la IA va ser de 2.76 punts sobre 5 ($sd = 0.7$), amb major puntuació dels professionals d'infermeria i dels qui van indicar utilitzar o no saber si utilitzaven la IA.

Knowledge and Perception of the Use of AI and its Implementation in the Field of Radiology: Cross-Sectional Study. Miró Catalina, Q., Femenia, J., Fuster-Casanovas, A., Marin-Gomez, F. X., Escalé-Besa, A., Solé-Casals, J., & Vidal-Alaball, J. (2023). *Journal of Medical Internet Research*, 25, e50728.

La revista *Journal of Medical Internet Research* (JMIR) té un factor d'impacte al 2023 de 7.4, ocupa la posició Q1 segons *Journal Impact Factor* (Clarivate, 2023) i es troba indexada a *Medline*, *Pubmed/PMC*, *Scopus*, *Psycinfo*, *SCIE*, *JCR*, *EBSECO Essentials*, *DOAJ*, *GoOA* i altres. És la revista insígnia de *JMIR Publications*. Líder mundial en volum, qualitat i visibilitat en l'àmbit de la salut digital. Publica articles sobre tecnologies emergents, dispositius mèdics, aplicacions, enginyeria, telemedicina i aplicacions informàtiques per l'educació del pacient, la prevenció, la salut de la població i l'atenció clínica.

Objectiu: descriure la percepció i coneixement de la població sobre la IA i el seu ús com a eina sanitària i la seva aplicació en l'àmbit de la radiologia, a través d'una enquesta validada.

Resultats: es varen obtenir 379 respostes, amb una mitjana d'edat de 43.9 ($sd = 17.52$) anys i un 59.8% dones. A més, el 89.8% de les respostes van indicar que entenen el concepte de la IA. De les quatre dimensions analitzades, les puntuacions properes a 5 indicaven percepcions negatives i properes a 1 indicaven percepcions positives. Així doncs, *desconfiança i responsabilitat* va obtenir una puntuació mitjana de 3.37 ($sd = 0.53$), *interacció personal* de 4.37 ($sd = 0.60$), *eficiència* de 3.06 ($sd = 0.73$) i *estar informat* de 3.67 ($sd = 0.57$). En relació a la dimensió de *desconfiança i responsabilitat*, les dones, la població major de 65 anys, la població amb estudis universitaris i la població que indicava no entendre el concepte d'IA, van presentar més negativitat. A la dimensió *estar informat* es va observar que el grup

amb estudis universitaris van valorar més positivament l'accés a la informació i que els qui van indicar no entendre el concepte d'IA la van valorar més negativament.

Real-world testing of an artificial intelligence algorithm for the analysis of chest X-rays in primary care settings. Miró Catalina, Q., Vidal-Alaball, J., Fuster-Casanovas, A. Escale-Besa, A., Ruiz Comellas, A., Solé-Casals, J. *Scientific Reports* 14, 5199 (2024).

La revista *Scientific Reports* és d'accés obert i té un factor d'impacte, al 2023, de 4.6, ocupa la posició Q1 segons *Journal Impact Factor* (Clarivate, 2023) i es troba indexada a *Web of Science*, *PubMed*, *PubMed Central*, *Scopus*, *Dimensions*, *Google Scholar*, *DOAJ* i *SAO/NASA ADS*. Publica articles originals sobre totes les àrees de les ciències naturals, la psicologia, la medicina i la enginyeria. És la cinquena revista científica més citada a nivell mundial i rep una àmplia atenció en documents polítics i mitjans de comunicació.

Objectiu: realitzar una validació externa, en la pràctica clínica real, de l'algorisme d'IA *ChestEye* per la lectura de radiografies de tòrax respecte els diagnòstics del radiòleg, així com detectar possibles diagnòstics pels quals l'algorisme no hagi estat entrenat.

Resultats: la validació externa es va realitzar amb una mostra de 278 imatges i informes, sent el 51.8% d'aquestes imatges sense alteracions radiològiques segons l'informe del radiòleg. Analitzant la validesa de l'algorisme d'IA, la mitjana de l'*accuracy* va ser de 0.95 (IC 95% 0.92; 0.98), la mitjana de la sensibilitat va ser de 0.48 (IC 95% 0.30; 0.66) i la de l'especificitat de 0.98 (IC 95% 0.97; 0.99). Les patologies on l'algorisme va tenir més sensibilitat van ser els implants externs, d'abdomen superior i cardíaca i/o valvular. En canvi, les patologies on l'algorisme va tenir menys sensibilitat van ser mediastí, vasos i òssia.

1.2 Difusió dels resultats de la tesi a congressos o jornades

A més a més de les publicacions a revistes científiques, a continuació es mostren les presentacions dels resultats de la tesi que s'han portat a terme en diferents congressos o jornades:

Queralt Miró Catalina, Jordi Solé Casals i Josep Vidal Alaball. Validació i avaluació de l'impacte de la intel·ligència artificial per la lectura de radiografies de tòrax. [Jornada InnoBages](#) – Impuls de la R+D+I al Bages organitzada pel Consell Comarcal del Bages, 21 de desembre del 2022. Comunicació oral.

Queralt Miró Catalina, Aïna Fuster Casanovas, Josep Vidal Alaball i Jordi Solé Casals. La intel·ligència artificial; el futur aliat de la radiologia. [Jornades dels 25 anys de l'IDIAP Jordi Gol](#), 28 de juny del 2022. Presentació en format pòster.

Queralt Miró Catalina, Aïna Fuster Casanovas, Jordi Solé Casals i Josep Vidal Alaball. Validació d'un algorisme d'intel·ligència artificial per la lectura de plaques de tòrax. [10a Jornada Sanitària del Món Rural de Berga](#) organitzada pel Col·legi de Metges del Berguedà, 5 de maig del 2023. Presentació en format pòster.

Queralt Miró Catalina, Aïna Fuster Casanovas, Josep Vidal Alaball, Anna Escalé Besa, Francesc X Marín Gómez, Joaquim Femenia i Jordi Solé Casals. Coneixements i percepció dels professionals sanitaris d'atenció primària sobre l'ús de la intel·ligència artificial com a eina sanitària. [10a Jornada Sanitària del Món Rural de Berga](#) organitzada pel Col·legi de Metges del Berguedà, 5 de maig del 2023. Comunicació oral i premi al millor treball científic.

Queralt Miró Catalina, Josep Vidal Alaball i Jordi Solé Casals. Validació i avaluació de l'impacte de la intel·ligència artificial per la lectura de radiografies de tòrax. [III Jornada d'Innovació de l'IDIAP Jordi Gol](#), 31 d'octubre del 2023. Comunicació oral.

Queralt Miró Catalina, Josep Vidal Alaball i Jordi Solé Casals. Intel·ligència artificial com a eina de suport a la salut pública. Jornada de Recerca de la Facultat de Ciències, Tecnologia i

Enginyeries de la Universitat de Vic – Universitat Central de Catalunya, [One Health: from environment to individuals, 2023](#). Presentació en format pòster.

A més, es va participar a les edicions del 2023 i 2024 de la [Doctoral Training Week](#) organitzada per la Universitat de Vic – Universitat Central de Catalunya. Les sessions on es va assistir van ser:

Cursos de *Mindfulness*, Taller de tesis doctorals, *Graphic desing for presentations and scientific communication*, *Que impacto ?* i *Las nuevas princesas toman la palabra: traducciones feministas de los cuentos de hadas* a l'edició del 2023 i *Ética, posibilidades y hype en la Inteligencia Artificial: Un Compromiso Científico* a l'edició del 2024.

1.3 Difusió dels resultats de la tesi als mitjans de comunicació

També s'ha presentat la tesi a diferents mitjans de comunicació:

Regió 7 (4 d'agost de 2022). L'ICS Catalunya Central té els primers professionals que fan doctorats industrials. Recuperat de: <https://www.regio7.cat/catalunya-central/2022/08/04/lics-catalunya-central-els-primers-73068767.html>

Nació Digital Manresa (4 d'agost de 2022). L'ICS Catalunya Central té els primers professionals que fan doctorats industrials. Recuperat de:

<https://www.naciodigital.cat/manresa/noticia/104653/societat/ics-catalunya-central-tesis-salut-industrial>

El 9 Nou (31 d'agost de 2022). Algoritmes per a la salut. Recuperat de: <https://el9nou.cat/osona-ripolles/actualitat/algoritmes-per-a-la-salut/>

El Regió 7 (8 de setembre de 2022). L'ICS de la regió central estudia analitzar radiografies amb intel·ligència artificial. Recuperat de: <https://www.regio7.cat/bages/2022/09/08/l-ics-regio-central-estudia-75161371.html>

L'entrevista central: Queralt Miró i Catalina (23 de setembre de 2022). Delegacions territorials del Govern a Catalunya. Recuperat de:

<https://delegacionscatalunya.gencat.cat/ca/delegacions/catalunya-central/fem-territori/peces-centrals/entrevistes/67/index.html>

Institut Català de la Salut Catalunya Central (8 de maig de 2023). Gran representació de professionals de l'ICS premiats en la 10a edició de les Jornades Sanitàries al Món Rural de Berga.

Recuperat de:

<https://catcentral.cpd1.grupics.intranet/web/index.php?seccio=noticies&id=11494&innovem=1>

Institut Català de la Salut Catalunya Central (2 d'octubre de 2023). Professionals de l'ICS Catalunya Central són premiats en l'acte de la Professió Mèdica del CoMB a Osona. Recuperat de: <https://catcentral.cpd1.grupics.intranet/web/index.php?seccio=noticies&id=11831>

L'Apunt, revista digital de la UVic (5 d'octubre de 2023). La professora de la FEC Queralt Miró, guardonada al XXIV Premi Ajut Consultori Bayés en Ciències de la Salut. Recuperat de: <https://apunt.uvic.cat/la-professora-de-la-fec-queralt-miro-guardonada-al-xxiv-premi-ajut-consultori-bayes-en-ciencies-de>

Regió 7 (16 d'octubre de 2023). L'ICS Catalunya Central analitza els coneixements i l'impacte de la IA en professionals i ciutadans. Recuperat de: <https://www.regio7.cat/catalunya-central/2023/10/16/l-ics-catalunya-central-analitza-93394157.html>

Nació Digital Manresa (16 d'octubre de 2023). L'ICS Catalunya Central publica un estudi sobre l'impacte de la intel·ligència artificial a la sanitat. Recuperat de:

<https://www.naciodigital.cat/manresa/noticia/111858/estudis-intelligencia-artificial-sistema-sanitari-queralt-miro>

L'Apunt, revista digital de la UVic (23 de novembre de 2023). Un estudiant de Medicina participa en un estudi sobre l'ús de la IA com a eina sanitària. Recuperat de: <https://apunt.uvic.cat/un-estudiant-de-medicina-participa-en-un-estudi-sobre-lus-de-la-ia-com-a-eina-sanitaria>

1.4 Ajuts i premis obtinguts

Primer premi al projecte de tesi doctoral del XXIV ajut Consultori Bayés en Ciències de la Salut a l'acte de la professió mèdica organitzat per la Junta Comarcal d'Osona del Col·legi de Metges de Barcelona (CoMB) i l'Agrupació de Ciències Mèdiques d'Osona (ACMO), 29 de setembre del 2023.

Premi al millor treball científic de l'article *Coneixements i percepció dels professionals sanitaris d'atenció primària sobre l'ús de la intel·ligència artificial com a eina sanitària* a la 10a Jornada Sanitària del Món Rural de Berga organitzada pel Col·legi de Metges del Berguedà, 5 de maig de 2023.

2. INTRODUCCIÓ

1.1 Origen i evolució de la intel·ligència artificial

L'any 1950, Alan Turing, un dels considerats com el pare de la informàtica, va presentar un treball anomenat "*Computing Machinery and Intelligence*" on es va formular, per primera vegada, la següent pregunta: les màquines poden pensar? A partir d'aquí, i al llarg d'aquesta dècada, es va confirmar que l'ordinador era el mitjà més prometedor per desenvolupar una intel·ligència creada per l'home [1,2].

Tot i això, no va ser fins el 1956 que va ser utilitzat per primer cop el terme d'intel·ligència artificial (IA) quan John McCarthy, Marvin Minsky, Nathaniel Rochester i Claude Shannon van presentar una proposta per realitzar un estudi sobre IA a *Dartmouth College, New Hampshire*, amb l'objectiu de demostrar que tot aspecte de l'aprenentatge o qualsevol característica de la intel·ligència pot descriure's amb tant detall i precisió que una màquina pot ser capaç de simular-lo i replicar-lo [1]. John McCarthy, aleshores, va definir la IA com la ciència i enginyeria de la fabricació de màquines intel·ligents, especialment de programes informàtics intel·ligents [3]. Durant les dues dècades següents, es van desenvolupar eines d'IA exitoses com el programa *General Problem Solver*, l'any 1957, capaç de resoldre problemes senzills automàticament, o ELIZA, el 1964, una eina de processament del llenguatge natural per simular conversacions amb humans [2].

Anys després, el 1995, Stuart Russel i Peter Norvig van publicar el llibre *Artificial Intelligence: A Modern Approach (AIMA)* convertint-se avui en dia en un dels llibres de text de referència sobre la IA, traduït a 15 idiomes, adaptat pel seu ús a més de 1.500 escoles i universitats de 134 països o regions, i actualitzat amb una 4a edició publicada a l'abril de 2020 [4].

Segons Russel i Norvig, la IA és l'estudi d'agents intel·ligents que reben inputs de l'entorn i actuen. Cada agent és implementat a través d'una funció que assigna percepcions i accions, i cobreix diferents formes de representar aquestes funcions, com sistemes de producció, agents reactius, planificadors lògics, xarxes neuronals o sistemes de decisió teòrics [5].

A més, al seu llibre, Russel i Norvig separen les definicions de la IA en funció de la racionalitat i pensament davant l'actuació. Segons ells, es pot classificar d'acord amb:

- Algorismes que pensen com els humans: alguns investigadors tenen com a objectiu crear màquines que pensin com els humans per poder reproduir els resultats o processos del pensament humà.
- Algorismes que actuen com els humans: amb l'objectiu de crear màquines que actuïn com els humans posant èmfasi en l'acció de manera més "robòtica", i no en el pensament per arribar a aquesta acció.
- Algorismes que pensen racionalment: amb l'objectiu de crear màquines que simulin el pensament racional i lògic dels humans, per tal de percebre i raonar davant d'una situació, com per exemple els sistemes de recomanació d'experts.
- Per últim, algorismes que actuen racionalment: hi ha qui vol desenvolupar algorismes per poder actuar de manera òptima i prendre les millors decisions de manera racional, simulant el comportament humà, percebent l'entorn com un agent intel·ligent [5].

Recentment, alguns dels experts en IA la defineixen com el camp que estudia la síntesi i anàlisi dels agents computacionals que actuen de manera intel·ligent (Poole i Mackworth 2010), o bé com la capacitat d'un sistema per interpretar correctament dades externes, aprendre d'aquestes i utilitzar aquests aprenentatges per aconseguir objectius i tasques específiques a través d'una adaptació flexible (Kaplan y Haenlein 2019) [5].

A dia d'avui, segons la Real Acadèmia Espanyola (RAE), la IA és la disciplina científica que s'ocupa de crear programes informàtics que executen operacions comparables a les que realitza la ment humana, com l'aprenentatge o el raonament lògic, o segons l'Institut d'Estudis Catalans, la IA és un conjunt de tècniques i reflexions teòriques sobre la construcció de sistemes intel·ligents [6,7]. I, fins i tot, segons el *ChatGPT* la IA és un camp de la informàtica que té per objectiu la creació de sistemes i programes computacionals que poden realitzar tasques que requereixen de la intel·ligència humana com l'aprenentatge, la resolució de problemes, o la presa de decisions, entre d'altres [8].

Amb tot això, tot i que el concepte ha anat evolucionant i s'ha anat adaptant al llarg del temps, la idea no ha variat i ha esdevingut una realitat amb un impacte no imaginat en tots els aspectes de la societat. La IA té tanta capacitat, encara desconeguda, que es podria arribar a definir com tot allò que les màquines encara no saben fer, o dit d'una altra manera, tota la capacitat que tenen per fer coses noves. Ens trobem, doncs, immersos en una revolució informàtica o digital, una nova era on la IA es troba subjacent en la majoria d'activitats de la vida quotidiana.

1.2 Definició de la intel·ligència artificial

Intel·ligència significa la capacitat d'entendre, comprendre, resoldre problemes. Artificial significa fet per l'home, no natural, produït per l'enginy humà. Així doncs, intel·ligència artificial és la creació d'algorismes a través de la capacitat humana per tal que aquests puguin imitar accions i prendre decisions. Amb altres paraules, la IA estudia la ciència i enginyeria per crear algorismes que puguin desenvolupar coneixements aprenent de l'experiència, llegint i processant textos de llenguatge natural, raonant a través de coneixements adquirits i actuant racionalment [3].

Per tant, la IA és la interacció de la informàtica, la lògica, la biologia, la psicologia, la filosofia, les matemàtiques, l'economia, la lingüística i moltes d'altres disciplines, que ha aconseguit generar aplicacions que permeten el reconeixement de veu, el processament d'imatges i de llenguatge natural, els motors de cerca, els sistemes de recomanació, així com els robots intel·ligents [9], entre d'altres. En particular, una tipologia d'IA n'és la IA generativa, capaç de crear idees i continguts nous com converses, imatges o vídeos a partir d'una petició. Un gran exemple d'IA generativa és el *ChatGPT* o el sistema de generació d'imatges *Stable Diffusion* [10,11].

Els algorismes d'IA es nodreixen de grans quantitats de dades que anomenem *Big Data*, i que avui en dia i amb el fenomen de la Internet de les Coses (IoC), som capaços de generar, gestionar i emmagatzemar. El *Big Data* sovint es defineix com les "tres v": volum (grans

quantitats de dades), velocitat (gran velocitat d'accés i anàlisi) i varietat (heterogeneïtat de les dades dels subjectes i de tipologia de data), és a dir, son bases de dades de gran varietat, en grans volums i sovint complexes, generades a grans velocitats [12]. Al ser volums de dades tan elevats, cal utilitzar programari de processament de dades amb molta capacitat per poder treballar-les i així resoldre problemes que no és possible abordar amb programari tradicional. Per això, gràcies al gran volum de dades, a la capacitat dels nous programes, i als algorismes desenvolupats pels humans, la IA és una realitat.

Quan parlem de la IoC ens referim a un sistema tecnològic que permet que els objectes es connectin a Internet i entre ells, connectant gairebé tot el que fa l'individu i controlant tots els dispositius d'aquest des de qualsevol lloc. Per exemple, els telèfons intel·ligents, rellotges intel·ligents, electrodomèstics intel·ligents, televisions intel·ligents, automòbils, entre d'altres. La IoC està, per tant, constantment generant dades i dades sobre gairebé tots els aspectes del dia a dia de les persones.

En aquest context de grans volums de dades, recentment s'està parlant d'una nova eina revolucionària i prometedora capaç de descentralitzar les dades: la *blockchain*. Es tracta d'una eina d'emmagatzematge i gestió de les dades que ofereix més seguretat de l'actual, més protecció de les dades contra alteracions o frauds i assegura una transparència de les dades impossibilitant un canvi d'aquestes un cop han estat creades. A més, la tecnologia *blockchain* ofereix la possibilitat de la creació d'un ecosistema descentralitzat i segur per compartir dades de salut, de garantir la seguretat i transparència de dades en l'àmbit de la recerca mèdica, així com facilitar la firma digital i l'emmagatzematge de consentiments informats. No obstant això, tot i representar una eina molt prometedora, tot just es troba en els seus inicis i li queda un llarg camí per recórrer abans de ser una realitat i trencar les actuals barreres com son les regulacions i les lleis, la complexitat tècnica, els costos tecnològics i econòmics, la falta d'experiència i confiança i les consideracions ètiques [13].

1.2.1 Mètodes que utilitza la intel·ligència artificial

En l'àmbit de la IA, la manera per ensenyar o entrenar els algorismes és mitjançant l'aprenentatge, resumint les lleis i mètodes d'identificació d'objectes a través de models de *machine learning* o aprenentatge automàtic, aprenentatge profund o *deep learning*, o les xarxes neuronals o *neural networks* (Figura 1). Per exemple, si s'introdueixen diverses imatges de gats a un ordinador, aquest és capaç d'aprendre les característiques dels gats a través de models d'entrenament i reconèixer els gats d'altres imatges a través de les característiques apreses. És a dir, els humans donen la capacitat d'aprendre a l'ordinador sense haver estat explícitament programat per fer-ho [9].

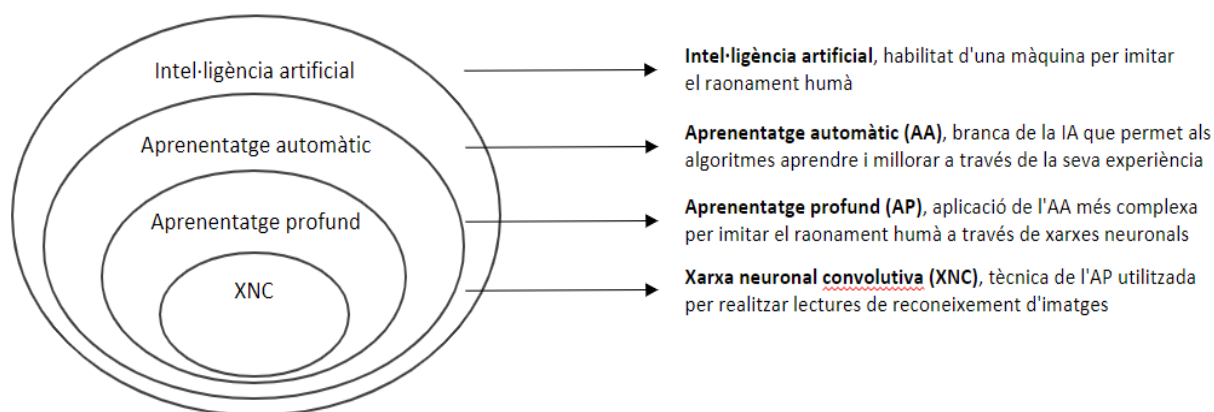


Figura 1. Mapa conceptual de les aplicacions i tècniques que utilitza la IA.

1.2.1.1 Aprenentatge automàtic

L'aprenentatge automàtic o *machine learning* és una branca de la IA que permet els ordinadors optimitzar processos, aprendre a través de grans bases de dades per tal de realitzar prediccions, prendre decisions i millorar automàticament a través de la seva experiència. Aquest aprenentatge es pot agrupar en quatre grups: aprenentatge automàtic supervisat, aprenentatge automàtic no supervisat, aprenentatge automàtic semi-supervisat i aprenentatge per reforç [9].

- L'aprenentatge automàtic supervisat és aquell que utilitza etiquetes per ser entrenat, és a dir, les dades d'entrada al model per ser entrenat ja es troben etiquetades amb el resultat correcte com si es tractés d'un supervisor ensenyant a una màquina a predir el resultat correctament. Alguns exemples en són el *Support Vector Machine* (SVM) o la discriminació lineal i sovint s'utilitzen per avaluar riscos, classificar imatges, filtrar *spam*, entre d'altres [14,15].
- L'aprenentatge automàtic no supervisat, en canvi, no utilitza etiquetes i també s'anomena *data mining*. En aquest cas, al no haver-hi etiquetes, els algorismes no s'entrenen de manera supervisada, sinó que les dades s'agrupen segons les seves característiques i permeten trobar patrons ocults dins el conjunt de dades. Els mètodes més utilitzats d'aprenentatge automàtic no supervisat són els algorismes d'agrupació, com el mètode de *clustering k-means*, o l'anàlisi de components principals.
- L'aprenentatge automàtic semi-supervisat apareix de la barreja entre el supervisat i el no supervisat, i els algorismes més utilitzats són basats en grafs i en màquines de vectors de suport semi-supervisats (S3VM). Sovint, aquest mètode s'utilitza quan poder obtenir les categories de totes les dades d'entrenament pot resultar molt costós. D'aquesta manera, en comptes d'etiquetar tot el conjunt de dades, s'etiqueta manualment una petita part, que s'utilitza per fer un primer entrenament i a continuació, s'aplica el model a tot el conjunt de dades sense etiquetar.
- Finalment, l'aprenentatge de reforç es basa en entrenar el model a base de recompenses. D'aquesta manera, l'aprenentatge de l'algorisme és a base d'assaig i error. A través de l'experiència, és capaç de millorar-se i adaptar-se a l'entorn. A diferència dels mètodes anteriors, en aquest cas no hi ha una resposta o categoria correcta sinó que l'algorisme ha de decidir què fer per realitzar una tasca concreta.

1.2.1.2 Aprenentatge profund

L'aprenentatge profund és una de les branques de l'aprenentatge automàtic basat en xarxes neuronals. Aquestes xarxes permeten treballar amb una gran quantitat de capes ocultes no lineals capaces de treballar amb estructures més complexes. Mentre que l'aprenentatge automàtic extreu les característiques al principi i és amb aquestes que entrena el model, l'aprenentatge profund treballa des de l'inici amb dades crues i les característiques les extreu dins del model a través de les capes inicials ocultes. Aquest mateix fet fa que els algorismes d'aprenentatge profund requereixin de més paràmetres per tal de preparar el sistema. Tot i no ser tan comú com l'aprenentatge automàtic, l'aprenentatge profund és molt útil en algunes aplicacions com per exemple en el desenvolupament dels cotxes autònoms o en la segmentació d'objectes en imatges i vídeos.

1.2.1.3 Xarxes neuronals convolucional

Una xarxa neuronal artificial és un model matemàtic inspirat en el funcionament i comportament biològic de les neurones. Per entendre bé el funcionament de les xarxes neuronals cal començar per la unitat més bàsica i simple, el perceptró. El funcionament del perceptró es mostra a la Figura 2 i es resumeix en: introducció dels valors d'entrada x_i (capa d'entrada), multiplicació de cada entrada x_i amb el seu pes corresponent w_i , suma de tots els productes anteriors, i aplicació d'una funció d'activació (normalment no lineal) sobre aquest resultat, per obtenir el que es coneix amb el nom de valor final (capa de sortida y) [16].

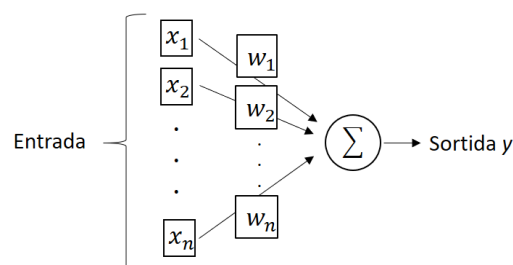


Figura 2. Funcionament del perceptró simple.

El perceptró només té la capacitat de separar o discriminar patrons en un hiperplà. En el cas de dades d'entrada de dues dimensions (2D), el perceptró separa a partir d'una recta que ve definida pels pesos de la xarxa. En el cas de dades d'entrada de tres dimensions (3D), el perceptró separa a partir d'un pla. Per poder generar fronteres (separacions) arbitràries, s'afegeixen capes ocultes entre mig (Figura 2 i Figura 3). Podem parlar de tres tipus diferents de capes en una xarxa neuronal [17]:

- Capa d'entrada: conjunt de dades d'entrada que s'utilitza per entrenar la xarxa neuronal. Tindrà tantes unitats com dimensions de les dades d'entrada.
- Capa oculta: conjunt de neurones que es troben entre la capa d'entrada i la de sortida. El nombre d'unitats pot ser qualsevol.
- Capa de sortida: conjunt de vectors o valor final. El nombre d'unitats està relacionat amb la dimensió de la sortida (dependrà de com es vol codificar el resultat de la xarxa neuronal).

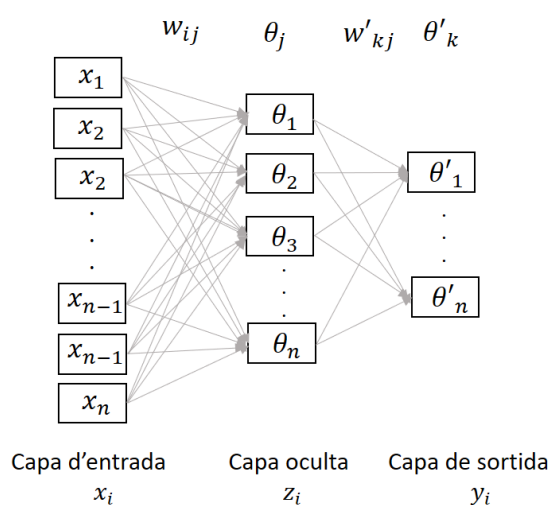


Figura 3. Funcionament del perceptró multicapa.

Una tipologia de les xarxes neuronals artificials de l'aprenentatge profund són les xarxes neuronals convolucionals (*Convolutional Neural Network, CNN*). Aquestes funcionen a través d'una capa d'entrada, una de sortida, i diverses capes ocultes d'aprenentatge i convolutives, i són utilitzades, entre d'altres, per realitzar lectures i reconeixement d'imatges. Una CNN

conté diverses capes per tal de fer la lectura d'un gran volum de dades i aprendre les característiques més rellevants d'aquestes. En el cas d'imatges com a dades d'entrada, les xarxes CNN *convolucionen* la lectura d'imatges per extreure'n informació. Una convolució és una operació matemàtica entre dues funcions (x i y) que produeix una tercera funció, normalment una modificació de x en funció de y . Un exemple d'algorisme de CNN són els sistemes de recomanació, de classificació, les anàlisis d'imatges mèdiques, o el processament del llenguatge natural, entre d'altres [18,19]. A la Figura 4 es pot veure el funcionament bàsic de les CNN:

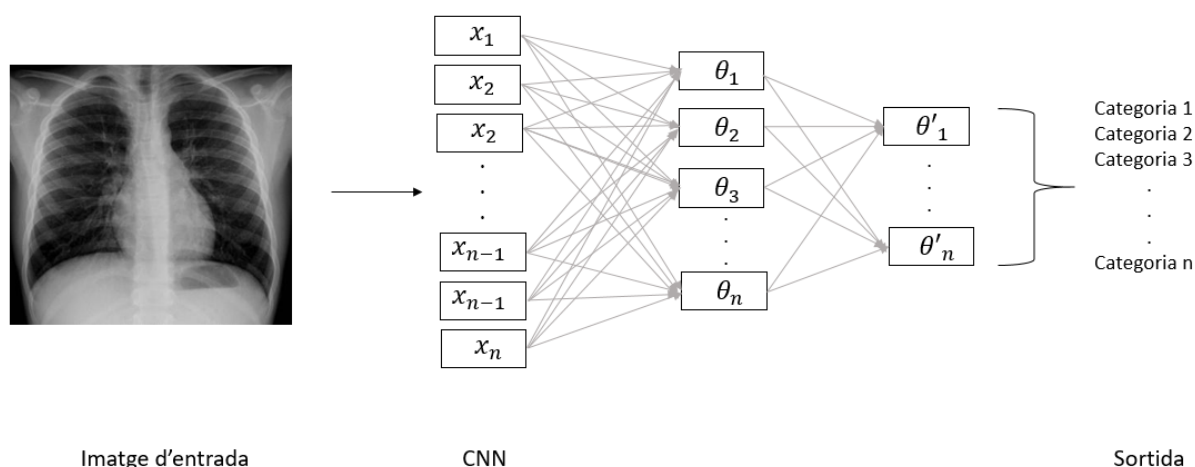


Figura 4. Funcionament de processament d'una imatge a través de les CNN.

1.3 Intel·ligència artificial en la societat

Els àmbits on la IA pot ser una ajuda i significar un avanç són molts, fins i tot en aquells on a dia d'avui no li trobem cabuda. Com l'aparició de la Internet o les xarxes socials, la IA impactarà a la vida diària de la mateixa manera o encara amb més força. La seva implementació pot fer la societat més eficient, reduir errors propis de la variabilitat humana i reduir despeses en recursos humans o derivades d'errors, entre d'altres [20].

Tot i això, una qüestió important és determinar quin paper exercirà la IA i, sobretot, com podran coexistir les aplicacions d'IA amb els humans. Quines decisions haurà de prendre la IA, quines els humans, i quines hauran de ser preses de manera conjunta? Tot i les demostracions de l'impacte positiu de les aplicacions d'IA, no totes són acceptades per la societat degut a un cert desconeixement, desconfiança i rebuig. Així doncs, si es vol preparar a la societat i elaborar estratègies d'acceptació de la IA, cal primer conèixer a fons quina és la percepció sobre la IA i quins són els coneixements que se'n té d'aquesta.

1.3.2 Percepcions i coneixements de la ciutadania

En l'actualitat encara hi ha pocs estudis que abordin el coneixement i la percepció de la població sobre la introducció de la IA en àmbits del seu dia a dia [21–23]. Tot i això, la poca literatura existent ha observat un gran interès de la ciutadania per la ciència i tecnologia, però poca familiarització amb el tema [24,25]. És important, doncs, empoderar la població i proporcionar-li els coneixements necessaris perquè estigui preparada i pugui entendre aquesta nova revolució digital. Per això, cal treballar la percepció i actituds socials enfront la ciència i tecnologia, la cultura científica-tècnica de la societat i l'interès de la ciutadania per la ciència [26].

Segons un estudi que va realitzar enquestes i entrevistes per avaluar l'impacte percebut de la societat sobre la IA i el que s'espera d'ella en un futur, existeix certa esperança en el seu impacte positiu en la capacitat de proporcionar resultats precisos i fiables en diferents àmbits. En el camp de l'educació facilitant i ampliant les metodologies d'estudi i aprenentatge, en el camp del transport creant noves formes de transport i conducció, en el camp de la salut, millorant directament la salut de la població, i en molts altres camps com en el camp de la presa de decisions o fins i tot en el camp de l'entreteniment [20].

No obstant això, segons l'àmbit d'aplicació de la IA, també hi ha certa preocupació sobre el seu impacte negatiu degut principalment a la desconfiança, als baixos nivells d'acceptació i a

l'escepticisme en relació a la seguretat de les dades i equitat dels sistemes. A més, a través d'alguns estudis s'ha observat certa desconfiança de la població sobre la capacitat dels governs en regular l'ús de la IA [27–29].

Recentment un dels temes més rellevants de l'aplicació de la IA no n'és la seva capacitat per millorar molts aspectes de la vida de la població sinó que n'és la seguretat, l'ètica i la vigilància. Així doncs, és important comptar amb una bona regulació per no perdre mai el control sobre la IA ni vulnerar els drets humans, i és necessari posar l'èmfasi en tenir un bon control a nivell intern governamental, i també a nivell extern, creant per exemple reguladors independents experts en la IA [30].

En resposta a les necessitat anteriors, el febrer del 2024 el parlament de la UE ha aprovat la primera llei que regula la IA per assegurar que els algorismes d'IA siguin segurs, transparents, no discriminants, respectuosos amb el medi ambient i supervisats sempre per les persones [31]. Després de ser presentada per primera vegada al 2021, la llei ha sigut aprovada al 2024 però no entrarà en vigor fins el 2026. Aquesta llei permetrà classificar les aplicacions d'IA segons el risc que representen, des de risc mínim fins a risc inacceptable.

1.4 Intel·ligència artificial en l'àmbit de la salut

De tots els àmbits on la societat mostra una bona acceptació i certa esperança en l'aplicació de la IA, en destaca especialment l'àmbit de la salut on es considera que hi pot ajudar amb la seva fiabilitat i certesa en la presa de decisions, millorant la confiança en el diagnòstic, l'eficiència en el flux de treball i permeten una atenció més personalitzada, entre moltes d'altres aplicacions [27,28].

Va ser a principis dels anys 70 quan es va començar a parlar de la IA aplicada a la salut, creant-se el 1975 el primer congrés d'IA en medicina per part de l'Institut Nacional de Salut, celebrat a la Universitat de Rutgers [32]. Un any després, es va presentar el primer programa d'IA sobre el glaucoma anomenat CASNET, i el programa MYCIN el qual, a partir de dades retrospectives

clíniques del pacient, podia proporcionar informació i recomanacions ajustades al pacient. No obstant això, no va ser fins anys després, de la mà de l'aprenentatge profund i de grans conjunts de dades, que va tenir lloc el creixement exponencial d'aplicacions d'IA aplicades a la salut fet que va començar a plantejar la pràctica clínica habitual implementada fins aleshores.

En efecte, un estudi publicat el 2020 va mostrar com el nombre d'articles publicats en revistes científiques sobre la IA en salut havia incrementat un 6% del 1995 al 2010, un 23% del 2011 al 2014 i un 43% del 2014 al 2019, sent els problemes de salut més comuns el càncer, la depressió, l'Alzheimer, les malalties cardiovasculars i la diabetis [33]. A més a més, un estudi del *Global Market Insights* de l'any 2022 va indicar que la inversió en l'ús de la IA en salut passaria de 760 milions de dòlars el 2016 a 10.000 milions de dòlars el 2024, i un altre estudi es va demostrar que el 2020 ja havia superat els 4.000 milions de dòlars. Aquest estudi també va indicar que s'espera un creixement anual de més del 30% entre els anys 2021 i 2027, arribant a una inversió de 34.500 milions de dòlars l'any 2027 [34–36].

No és estrany aquest augment en recerca i inversió destinada a la IA, ja que a dia d'avui la IA en l'àmbit de la salut pot facilitar tasques administratives i fluxos de treball, utilitzar la telemedicina a través de programes per realitzar assistència virtual, donar suport a la realització de diagnòstics (d'imatge o sense), o permetre una atenció més personalitzada i preventiva, entre moltes d'altres aplicacions, millorant directament i indirecta l'experiència i la salut de la població.

Actualment ja hi ha aplicacions d'IA que s'utilitzen a les consultes diàries dels professionals sanitaris. No obstant això, ens trobem encara als inicis del seu ús i només s'han explorat una petita part de les capacitats de la IA. És per això que, en molts dels estudis realitzats, s'afirma que encara hi ha camí per recórrer realitzant més validacions amb dades independents i desenvolupant i ajustant els models a les necessitats reals de les consultes i de la població [32].

1.4.1 Percepcions i coneixements dels usuaris i professionals sanitaris

Diferents estudis on s'avalua les percepcions i coneixements de la població respecte la IA en l'àmbit sanitari ja han demostrat la necessitat d'ajustar la IA als usuaris, però el volum d'estudis on es demana a la població la seva percepció encara és baix. Fins al moment, la percepció de la població respecte la IA és positiva: la població creu que la IA en l'àmbit sanitari pot oferir solucions més personalitzades, més fàcils d'interpretar, més acurades i eficients, i fins i tot, donar la possibilitat de demanar una segona opinió, a part de la del professional de referència, basada en la IA i disponible les vint-i-quatre hores del dia [37–41].

Tot i això, també hi ha cert desconeixement en el funcionament intern de la IA i una preocupació en els problemes ètics, jurídics, de seguretat i de protecció de les dades que pot implicar. Un dels punts negatius generalitzats que troba la societat és que la seva implementació pugui afectar negativament la comunicació entre les persones, així com en la necessitat d'oferir solucions personalitzades i adaptades a totes les edats independentment del nivell informàtic que tinguin. A més, la població té clar que la IA aplicada a la salut no ha d'oferir diagnòstics definitius ni donar obligacions. Es tracta, doncs, d'una eina de suport al professional, per donar una altra visió i per oferir recomanacions als usuaris [37,39–41].

Segons el punt de vista dels professionals sanitaris, la majoria d'articles indiquen que són plenament conscients de la seva imminent aplicació i de l'afectació positiva que pot tenir en el seu dia a dia laboral i en la salut de la població. Actualment, el percentatge de professionals que utilitzen la IA en el seu dia a dia de la consulta és molt baix, però tot i així reconeixen l'impacte positiu en la millora que pot tenir la precisió diagnòstica, l'atenció sanitària, l'augment d'eficiència i productivitat, l'estalvi en despeses sanitàries, el millor seguiment del pacient, la detecció precoç de malalties, la millora en la prevenció i la reducció d'errors mèdics [42,43].

De totes maneres, hi ha altres articles que no deixen tan clara l'opinió positiva dels professionals respecte la IA indicant en alguns casos cert desconeixement sobre la IA així com por i dubtes sobre la possibilitat de ser reemplaçats per aquesta [44,45]. També hi ha

preocupació sobre la implementació de la IA degut al desconeixement intern del funcionament dels algorismes d'IA, perquè sovint aquests són percebuts com una caixa negra i no conèixer com funcionen genera rebuig en els seus resultats [46–48]. A més, hi ha preocupació respecte les consideracions ètiques, la pèrdua de control de la gestió i la responsabilitat del professional al utilitzar eines d'IA [42].

Per tant, la literatura actual indica una falta de formació i preparació per part dels professionals. La majoria d'aquests considera important realitzar formacions sobre IA, el seu funcionament, la seva implementació, el tractament de les dades que fa, entre d'altres, i resulta important oferir aquestes formacions abans de la implementació de la IA per poder brindar-los seguretat i confiança [43].

Així mateix, la majoria d'especialistes de la medicina consideren que els sistemes no es troben actualment preparats per la implementació de la IA i tampoc tenen capacitat per fer-ne validacions [38,45–48]. Tot i haver-hi aplicacions d'IA amb resultats molt prometedors en les validacions realitzades pels propis desenvolupadors, adaptar aquestes aplicacions a un producte que es pugui utilitzar en el dia a dia de les consultes és complicat, ja que cal assegurar-se que els resultats són realment robustos amb dades d'entorns reals i diferents a les d'entrenament.

Així doncs, i segons la literatura actual, a part de seguir entrenant la capacitat de la IA, desenvolupant nous models i ajustant-los a les necessitats dels professionals i usuaris, cal posar especial èmfasi en dos aspectes importants; d'una banda insistir en la necessitat de controlar els límits ètics, de protecció de dades, de transparència dels models i de governança de les dades i dels algorismes d'IA. D'altra banda, centrar-se en la percepció i coneixements dels usuaris i sobretot dels professionals sanitaris per unificar forces, treballar conjuntament totes les parts implicades per incrementar l'acceptabilitat de la IA i ajustar-la a les necessitats reals de la societat.

1.5 Radiologia

En el context de la IA en salut, una de les especialitats on hi ha major aplicacions actuals i un ventall d'aplicacions futures més gran és en la radiologia a través de l'aprenentatge profund i les CNN. La radiologia és una especialitat mèdica basada, principalment, en el diagnòstic i tractament de les malalties a través de la imatge. Es troba present en aproximadament el 10% de les consultes ambulatories, en el 50% dels pacients atesos a urgències i en el 70% dels pacients hospitalitzats [49]. Així i tot, la Organització Mundial de la Salut (OMS) estima que dos terços de la població mundial no té accés a serveis de radiologia [50].

El 2017, la *British Medical Journal* (BMJ) ja anunciava una important falta de radiòlegs per poder interpretar tota la demana d'imatges del Regne Unit. Segons l'informe del cens de personal de radiologia clínica del Regne Unit del *Royal College of Radiologists*, l'especialitat de radiologia va ser la segona especialitat amb menor proporció de residents per cada especialista respecte altres especialitats, i un de cada deu llocs de radiòlegs del Regne Unit (8.5%) van quedar vacants el 2016 [51,52]. Aquesta falta de radiòlegs també es va veure a l'informe publicat tres anys després, el 2019 [53] i a més, aquesta manca de radiòlegs es va accentuar durant el període de la COVID-19, moment en que hi havia una enorme demanda per interpretar radiografies de tòrax però no prou professionals per fer-ho, especialment en zones amb pocs recursos i en zones rurals [54–56].

D'altra banda, a part de la falta de radiòlegs, els informes radiològics com tota tasca desenvolupada pels humans, poden contenir errors o diversitat d'interpretacions segons el professional avaluador. L'any 1899 ja hi havia estudis que descrivien possibles errors en el diagnòstic degut a factors externs com la llum de la sala o la qualitat de la imatge [57]. A dia d'avui, el terme error en l'àmbit sanitari s'ha definit com un diagnòstic realitzat amb retard involuntari, com un diagnòstic erroni, un diagnòstic no identificat, o com la incertesa sobre la correcta interpretació [58,59].

En la radiologia, existeixen diverses classificacions segons la tipologia d'errors [59,60]. De forma resumida, els agrupen segons error degut a un fals positiu, error d'interpretació i/o

raonament, error per desconeixement o complexitat de la patologia, error en la comunicació entre professionals, error degut a una mala tècnica d'imatge, error per no visualitzar les imatges prèvies si es tracta d'un seguiment o una bona lectura de l'historial i sobretot l'error per no poder disposar de temps i espai necessari per fer una correcta interpretació. En especial, augmenta el risc d'aquests errors quan es tracta de tornos de guàrdies amb major volum d'imatges per interpretar, sovint de major complexitat, i amb una fatiga d'hores treballades acumulada [61–65].

Hi ha diversos estudis en l'àmbit de la radiologia, que descriuen la quantitat d'errors esmentats anteriorment. N'és un bon exemple un estudi realitzat als Estats Units fa gairebé deu anys on es va observar més de mil errors al revisar 600 casos complexes que havien obtingut un diagnòstic amb retard o un diagnòstic erroni, amb una mitjana de dos-cents cinquanta dies des de la primera interpretació errònia fins al diagnòstic correcte. Els errors es van classificar en diferents categories, segons si es tractava de falta d'informació de la història clínica, en errors en les interpretacions, en falta de comunicació entre professionals, errors o limitacions en la tècnica d'exploració, errors per falta de localització de la alteració o errors deguts a casos molt complexes [66].

Per aquests motius, des de fa anys s'està desenvolupant aplicacions d'IA de suport a la interpretació d'imatges radiològiques per poder minimitzar l'impacte de la falta de radiòlegs i el nombre d'errors, ajudant a la presa de decisions, en la fiabilitat i eficiència del diagnòstic així com en l'automatització de processos manuals per tal d'optimitzar millor el seu temps. A més, la IA ofereix la possibilitat d'extreure informació de les imatges que l'ull humà és incapaç de percebre, analitzar-les de manera massiva i combinar-les amb altres informacions rellevants del pacient per poder crear prediccions d'evolució d'aquest o deteccions més primerenques de malalties.

En l'àmbit de la radiologia, fa més de 60 anys que es va introduir el sistema de diagnòstic assistit per computador (CAD, *computer assisted diagnosis*) pel metge Gwilym S. Lowdick, i des de fa anys es treballa amb aquest sistema en la detecció de càncer de mama, nòduls al

pulmó o pòlips de colon, entre d'altres [67]. Tot i això, aquest sistema no és fàcil de desenvolupar i no té la capacitat de la IA de tenir una millora i entrenament constant.

1.5.1 Radiografia de tòrax

Una de les tècniques d'imatge més utilitzada en l'àmbit de la radiologia és la radiografia (Rx), una forma de radiació electromagnètica d'ona curta que travessa la matèria i és utilitzada per generar imatges de l'interior del cos a través del contrast de la radiació i no de la llum. Aquesta prova té l'avantatge de ser ràpida de realitzar i de requerir nivells molt baixos de radiació al realitzar-se (0.2 milisieverts o mSv) [68–70].

De totes les radiografies, un gran percentatge en són les radiografies de tòrax, una prova en format d'imatge que permet veure el cor, pulmons, vasos sanguinis, vies respiratòries, ossos del tòrax i la columna vertebral, i que sol ser utilitzada per detectar patologies pulmonars i cardiovasculars [71–74].

Hi ha diverses classes de radiografies de tòrax segons la seva projecció. Les dues projeccions frontals més utilitzades són l'anteroposterior (AP), on els rajos travessen el tòrax de la part anterior a la posterior i la posteroanterior (PA), on els rajos travessen el tòrax de la part posterior a la anterior, deixant totes dues una imatge frontal del tòrax. La projecció PA és més comuna, i la projecció AP sol utilitzar-se quan es tracta d'un pacient estirat que no es pot incorporar o quan la imatge es fa fora del departament utilitzant unitats mòbils de rajos. També hi ha la projecció lateral i altres projeccions menys utilitzades [75,76].

L'avaluació de les radiografies de tòrax és una tasca que desenvolupa majoritàriament el radiòleg, encara que també poden fer-la altres especialistes com els metges d'urgències, de família, cardíologs, pneumòlegs, entre d'altres, i que requereix una observació detallada i un coneixement apropiat. A més, les radiografies de tòrax representen una gran càrrega de treball dels serveis sanitaris, tot i que a alguns països hi ha una important falta de radiòlegs per interpretar-les i els errors de detecció en la seva interpretació, per diferents motius, són força freqüents [77,78].

1.5.2 Aplicació de la intel·ligència artificial a la radiologia

Considerant la combinació entre el gran volum de radiografies de tòrax, fet que implica una gran quantitat d'informació en format d'imatges, el reduït nombre de professionals per interpretar-les, i l'evolució de la informàtica, es fa molt atractiva la implementació d'aplicacions d'IA en aquesta disciplina, i explica la gran quantitat d'aplicacions desenvolupades i en procés de desenvolupament per la lectura d'aquestes imatges. A més, la pandèmia ocasionada per la COVID-19 n'ha sigut un catalitzador [79].

De totes maneres, al ser conceptes relativament nous, existeix certa preocupació sobre la possibilitat que la IA acabi substituint els radiòlegs i professionals del camp de la radiologia donada l'aparició d'algorismes capaços de desenvolupar tasques realitzades fins l'actualitat, per l'home. En el seu moment, també hi va haver qui va dubtar que la calculadora acabaria substituint l'enginyer. Res més lluny de la realitat, avui en dia som capaços de veure que la calculadora, o la tecnologia en general, no té per objectiu substituir l'humà, sinó fer-lo més eficient. Les màquines, per molt capaces de simular tasques pròpiament realitzades per humans, sempre necessitaran professionals per ser creades, entrenades, validades i utilitzades.

Per això, la gran majoria de professionals no considera que la IA pugui substituir els radiòlegs o altres professionals sanitaris, ja que les propostes d'interpretacions de la IA sempre hauran de ser supervisades per un especialista, especialment en casos complexos, i a més, aquests desenvolupen altres tasques que la IA no és capaç de fer com la radiologia intervencionista, la valoració de pacients a consulta externa, la creació de tractament i el seguiment de pacients, entre d'altres [80,81]. Es tracta, doncs, d'una eina que farà més eficient els professionals de la radiologia i d'altres especialitats mèdiques, ajudant-los en la presa de decisions i donant més temps per poder-se dedicar a tasques complexes si poden evitar, per exemple, analitzar casos sense alteracions descartats ja prèviament per la IA.

A més, les radiografies de tòrax sovint han de ser interpretades per professionals de la medicina d'altres especialitats a la radiologia. El diagnòstic radiològic també forma part de les

competències clíniques de la medicina familiar i comunitària, en la qual és habitual a la pràctica clínica dels seus professionals la interpretació de radiografies de tòrax malgrat el seu menor grau d'expertesa en relació als radiòlegs [82]. Així mateix, l'especialitat de medicina familiar i comunitària representa el primer contacte amb el sistema sanitari i això implica veure una gran varietat d'afeccions mèdiques; des d'afeccions més lleus a les més greus i urgents que poden arribar a representar el 5% dels casos d'atenció primària [83]. Per aquets motiu, disposar d'aplicacions d'IA com a eines de suport a la interpretació d'aquestes imatges pot resultar de gran ajuda en altres especialitats com la medicina familiar i comunitària.

1.6 Aplicacions actuals d'intel·ligència artificial per a la lectura de radiografies de tòrax

En l'àmbit de la radiografia de tòrax, en els darrers anys s'han desenvolupat aplicacions d'IA a través de CNN capaces de determinar si una imatge presenta alteracions radiològiques o no, i capaces de detectar moltes de les patologies que es poden veure a través d'una radiografia de tòrax. El 2021, hi havia 100 productes d'IA en l'àmbit de la radiologia amb la marca CE de 54 proveïdors diferents. No obstant això, només 36 d'aquests productes presentaven evidències de la seva eficàcia a través d'avaluacions de la precisió diagnòstica, la meitat d'aquests per l'anàlisi de radiografies de tòrax, i només el 18% presentava una validació per demostrar el seu impacte en el pensament, la salut del pacient i els costos econòmics [84–86].

Alguns dels algorismes desenvolupats fins al moment tenen com a objectiu el triatge de les radiografies realitzades, classificant-les segons si presenten o no alteracions o fins i tot segons el seu grau de gravetat [87,88]. D'altres algorismes s'han centrat amb algunes patologies en concret, com per exemple en la tuberculosi [89–93], en nòduls pulmonars malignes [89,94–102], en insuficiència cardíaca [103], en pneumotòrax [104], en malalties intersticials [105,106] o en pneumònia [107,108]. Fins i tot, s'han desenvolupat algorismes capaços de

detectar més d'una patologia en concret i capaços de pintar sobre la imatge la localització de la patologia [89,109,110].

Amb l'ús d'aquestes aplicacions es millora el rendiment dels professionals i es redueix el temps de lectura de les imatges per part dels professionals. S'ha vist que es pot reduir fins a quatre vegades el temps d'espera per l'informe d'una radiografia crítica i fins a dues vegades per una radiografia urgent [87,111].

Tot i això, considerant una validació externa com aquella que compara, amb dades independents a l'entrenament, els resultats obtinguts amb l'algorisme i els resultats obtinguts a través dels professionals, només un 6% dels articles publicats sobre algorismes d'aprenentatge profund en radiologia inclouen una validació externa [112]. Per tant, la majoria d'aplicacions d'IA per la lectura de radiografies de tòrax no presenten validacions externes, ni tampoc presenten validacions amb dades prospectives d'entorns clínics reals. Es tracta de validacions internes a través de dades retrospectives, sovint amb criteris de presència o absència de la malaltia, mal definits degut a que les dades havien estat recollides amb altres objectius [113,114].

En l'àmbit de l'atenció primària, hi ha pocs algorismes d'IA per la lectura d'imatges de tòrax amb validacions externes. L'estudi més recent presenta una validació d'un algorisme capaç de detectar nòduls pulmonars validat amb 5.722 imatges obtingudes de centres d'atenció primària del Regne Unit. Per tal de fer la validació, es compara el diagnòstic de l'algorisme amb el diagnòstic del radiòleg i el diagnòstic d'un equip de professionals sanitaris multidisciplinari. A través de la comparació de l'algorisme, es va obtenir una sensibilitat del 54.5% (IC 95%: 44.2%-64.4%) i una especificitat del 83.2% (IC 95%: 82.2%-84.1%) en comparació amb el radiòleg i una sensibilitat del 60.9% (IC 95%: 50.1%-70.9%) i una especificitat del 83.3% (IC 95%: 82.3%-84.2%) en comparació amb l'equip de professionals multidisciplinari. L'estudi conclou que hi ha una falta de generalització dels resultats de la validació interna i de l'entrenament i per tant, indica que hi ha una necessitat de validar els algorismes amb dades reals i representatives [115].

Una altre estudi molt recent va analitzar 3.047 imatges de tòrax d'atenció primària per validar externament un algorisme d'IA capaç de classificar les imatges en clínicament significatives (pneumònia, tuberculosi o càncer de pulmó), insignificants i normals. El valor de l'AUC en la lectura de les imatges sense l'ajuda de l'algorisme va variar entre 0.534-0.676 i amb l'ajuda de l'algorisme entre 0.571-0.688. També va observar que disminuïa significativament el temps de lectura de les imatges amb l'ajuda de l'algorisme entre 2 i 10 segons. No obstant això, tot i millorar l'eficiència dels professionals amb menys experiència, també va concloure obtenir uns resultats de validació externa pitjors als esperats segons les validacions internes i resultats d'entrenament [116].

Per últim, un estudi del 2022 va realitzar una validació externa d'un algorisme d'IA capaç de detectar 11 patologies a través de 1.960 imatges d'atenció primària al Regne Unit, comparant el diagnòstic del radiòleg amb el diagnòstic de l'algorisme. L'AUC que va obtenir l'algorisme va ser entre 0.881-0.998, els valors d'*accuracy* entre 0.754 i 0.986, els valors de sensibilitat entre 0.550 i 1 i els valors d'especificitat entre 0.674 i 0.996. L'estudi conclou que l'algorisme pot ser de gran ajuda per fer el triatge entre les imatges normals i les imatges amb alteracions majors [117].

1.6.1 Oxipit

Oxipit és una empresa Lituana, fundada l'agost de 2017, destinada a la creació d'eines d'IA per la lectura d'imatges mèdiques. En l'actualitat, Oxipit compta amb nou professionals i és una de les empreses líders en la lectura d'imatges mèdiques mitjançant IA. Ha desenvolupat o es troba actualment desenvolupant algorismes per a la lectura d'imatges de tòrax, de mamografies, de radiografies múscul-esquelètiques i de tomografia computada (TC) [118,119].

Tot i treballar en diferents àmbits d'imatge radiològica, la seva principal aplicació és un algorisme automàtic i autònom d'IA per la lectura de les radiografies de tòrax entrenat amb més de 300.000 imatges, disponible a través d'una plataforma web anomenada *ChestEye*. El

servei d'imatge *ChestEye* va néixer l'any 2017 i el març de 2022 comptava amb el certificat com a dispositiu mèdic de classe II al Registre Australià de productes terapèutics i també amb el certificat del marcatge CE IIa [118,119].

El funcionament de la plataforma web *ChestEye* és el següent: el professional insereix a l'aplicació la radiografia de tòrax anteroposterior en format DICOM, l'aplicació fa la lectura de la radiografia i retorna l'informe automàticament amb capacitat per detectar 75 patologies, les quals cobreixen el 90% dels diagnòstics, així com un mapa de calor per mostrar les localitzacions de les troballes (Figura 5). D'aquesta manera, *ChestEye* dona suport amb el diagnòstic i podria servir de triatge permetent que els professionals només hagin d'analitzar les radiografies més rellevants [118,119].

ChestEye ha estat desenvolupat mitjançant xarxes neuronals que han sigut entrenades a través de processos iteratius, amb grans quantitats d'imatges, per tal que l'algorisme pugui aprendre automàticament a partir de les imatges llegides. Es tracta d'una eina de suport molt prometedora que encara no ha estat validada externament amb dades reals.

Figura 5. Exemple d'ús de l'aplicació per a la lectura d'imatge.

OXIPIT About English

Priority **HIGH**

Findings


- An ill-defined patchy opacity is present in the left upper lung field.
- A well defined oval/rounded shape pulmonary nodule is seen in the left upper lung field (102 mm).
- A pulmonary cavity is present in the left upper lung field, possibly related to tuberculosis infection.
- There is air on both sides seen compressing the underlying lung parenchyma.
- Normal hilar vascular markings.
- The heart is normal in size.
- The mediastinum is within normal limits.

Impression

- Left upper lung field consolidation.
- Nodule in the left upper lung field (102 mm).
- Cavitary tuberculosis.
- Bilateral pneumothorax.

Heatmaps

- Consolidation
- Pneumothorax
- Tuberculosis
- Pulmonary Cavity
- Nodule



Zoom: 0.701
WWWC: 255/128

Invert
Flip Horizontally
Reset Image

Search for similar cases

Caution - INVESTIGATIONAL DEVICE, LIMITED BY FEDERAL (OR UNITED STATES) LAW TO INVESTIGATIONAL USE

OXIPIT About English

Priority **HIGH**

Findings

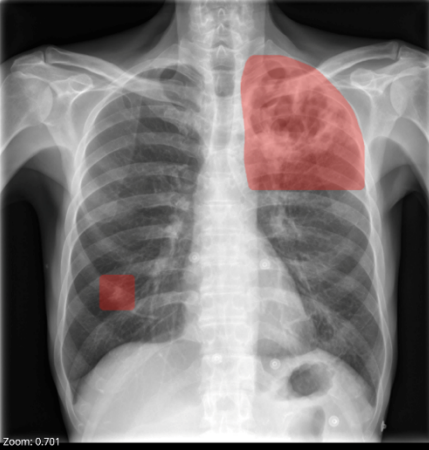
- An ill-defined patchy opacity is present in the left upper lung field.
- A well defined oval/rounded shape pulmonary nodule is seen in the left upper lung field (102 mm).
- A pulmonary cavity is present in the left upper lung field, possibly related to tuberculosis infection.
- There is air on both sides seen compressing the underlying lung parenchyma.
- Normal hilar vascular markings.
- The heart is normal in size.
- The mediastinum is within normal limits.

Impression

- Left upper lung field consolidation.
- Nodule in the left upper lung field (102 mm).
- Cavitary tuberculosis.
- Bilateral pneumothorax.

Heatmaps

- Consolidation
- Pneumothorax
- Tuberculosis
- Pulmonary Cavity
- Nodule



Zoom: 0.701
WWWC: 255/128

Invert
Flip Horizontally
Reset Image

Search for similar cases

Caution - INVESTIGATIONAL DEVICE, LIMITED BY FEDERAL (OR UNITED STATES) LAW TO INVESTIGATIONAL USE

OXIPIT About English

Priority **HIGH**

Findings

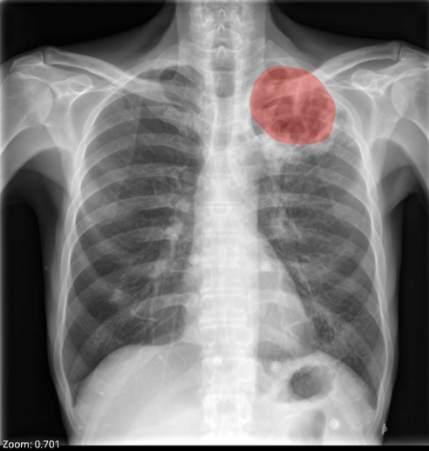
- An ill-defined patchy opacity is present in the left upper lung field.
- A well defined oval/rounded shape pulmonary nodule is seen in the left upper lung field (102 mm).
- A pulmonary cavity is present in the left upper lung field, possibly related to tuberculosis infection.
- There is air on both sides seen compressing the underlying lung parenchyma.
- Normal hilar vascular markings.
- The heart is normal in size.
- The mediastinum is within normal limits.

Impression

- Left upper lung field consolidation.
- Nodule in the left upper lung field (102 mm).
- Cavitary tuberculosis.
- Bilateral pneumothorax.

Heatmaps

- Consolidation
- Pneumothorax
- Tuberculosis
- Pulmonary Cavity
- Nodule



Zoom: 0.701
WWWC: 255/128

Invert
Flip Horizontally
Reset Image

Search for similar cases

Caution - INVESTIGATIONAL DEVICE, LIMITED BY FEDERAL (OR UNITED STATES) LAW TO INVESTIGATIONAL USE

1.7 Regulació i legislació de productes i dispositius d'intel·ligència artificial

Des de ja fa anys, la Unió Europea està treballant en el desenvolupament d'un marc on quedin reflectits tots els temes importants sobre la IA, des dels seus potencials usos fins al compliment legal de la seguretat i l'ètica. A l'abril de 2018, es va publicar una Estratègia Europea de Dades [120] per abordar la IA i establir mesures polítiques.

Juntament amb l'Estratègia Europea del 2018, es va crear un grup d'experts d'alt nivell que va publicar algunes indicacions per una IA fiable a través d'una comunicació on es parlava de l'acció i supervisió humana, solidesa tècnica i de seguretat, gestió de la privacitat i les dades, transparència, diversitat no discriminació i equitat, benestar social i mediambiental i rendició de comptes [121].

El 2020, a Brussel·les, es va publicar el Llibre Blanc sobre la IA: un enfoc europeu orientat a la excel·lència i la confiança [122]. En aquest llibre es recollia informació sobre la IA a través d'un enfoc europeu coordinat, amb l'objectiu de treballar en les implicacions ètiques i humanes de la IA i analitzar possibles millores d'ús de les macro-dades en la innovació.

Al Llibre Blanc sobre la IA també es parlava d'un futur marc regulador de la UE i d'altres requisits que cal que les aplicacions d'IA respectin en relació a les dades d'entrenament, a la conservació del registre de dades, a la transparència d'informació, la solidesa i exactitud i la supervisió humana. A més, també exposava que hi hauria d'haver controls en els assajos d'aquestes aplicacions, en les certificacions, en els algorismes i fins i tot en els conjunts de dades utilitzats en la fase de desenvolupament [122].

Tot i això, el 2021 un conjunt d'eurodiputats va reclamar més garanties en el compliment legal de la IA demanant aturar algunes aplicacions ja implementades d'IA. Exposaven la necessitat real de disposar d'un marc regulador clar a nivell europeu per generar confiança, respectar la protecció de dades i accelerar la seva acceptació. No ha sigut fins a principis del 2024 que el parlament de la UE ha aprovat la primera llei que regula la IA per assegurar que els algorismes d'IA siguin segurs, transparents, no discriminants, respectuosos amb el medi ambient i

supervisats sempre per les persones. Tot i això, la llei ha sigut aprovada el 2024 però no entrarà en vigor fins el 2026 [123].

Aquesta llei permetrà classificar les aplicacions d'IA segons el risc que representen, des de risc mínim fins a risc inacceptable. Les aplicacions de risc inacceptable estaran prohibides, per exemple els sistemes de puntuació social. Les aplicacions d'alt risc i risc limitat estaran regulades, mentre que les aplicacions de risc mínimes no estaran regulades com per exemple els sistemes de recomanació o videojocs. A més, la majoria de les obligacions definides recauen sobre els desenvolupadors de les aplicacions d'IA on tots han de proporcionar una documentació tècnica, instruccions d'ús, complir amb la directiva de drets d'autor i publicar un resum sobre els continguts utilitzats per al desenvolupament i entrenament dels algorismes. Fins i tot avaluacions dels algorismes i notificacions d'incidents en les aplicacions que presentin un risc sistèmic [123].

A nivell autonòmic, el 2023 s'ha creat la primera associació que treballa per defensar els interessos de la població respecte l'impacte de la IA assegurant un accés lliure a aquesta i un desenvolupament ètic. Aquesta associació s'anomena CIVICAI i està formada per 250 professionals i 136 fundadors [124]. L'objectiu de CIVICAI és sensibilitzar a la societat i promoure un debat crític i participatiu per tal de conscienciar-la de l'impacte que tenen els avenços constants de la IA així com d'implicar-la a la seva regulació i desenvolupament. Segons l'associació, la IA ha de ser de tots i per a tots, i per tant, ha d'estar al servei de les persones i de la societat de manera equitativa i responsable.

D'altra banda, a part de centrar els esforços en una regulació de la IA, també s'indica en la literatura la necessitat de crear requisits per al desenvolupament i prova dels algorismes d'IA, permetent una regulació estable tot i la pròpia evolució dels aspectes tècnics dels algorismes, i fins i tot una espècie de "codi moral" per reforçar la responsabilitat dels enginyers desenvolupadors d'algorismes o empreses desenvolupadores [2].

Adicionalment, i com ja s'ha mencionat anteriorment, actualment la capacitat de generació i emmagatzematge de les dades és molt àmplia perquè les dades sanitàries poden provenir de diverses fonts d'informació: historials mèdics, reclamacions d'assegurances, dispositius

wearables i xarxes socials, entre d'altres. Tenir més informació significa que si hi ha un filtratge o error en la privacitat de les dades, el volum de persones implicades serà més gran i per tant el dany serà major.

Als Estats Units, la privacitat de les dades sanitàries es tracta de manera diferent en funció de com s'han creat i qui les custodia [125–127]. En canvi, el Reglament General de Protecció de Dades (GDPR) de la Unió Europea estableix un únic règim definit per les dades sanitàries a través d'una categoria de “dades relatives a la salut” per referir-se a “dades personals relacionades amb la salut física o mental d'una persona física, inclosa la prestació de serveis d'assistència sanitària, que revelin informació sobre el seu estat de salut”.

En l'àmbit dels dispositius sanitaris que contenen IA, les directrius anteriors es mantenen i apareixen noves regulacions que cal complir. Per tal de poder utilitzar i comercialitzar un dispositiu sanitari generat amb IA cal que aquest disposi de la Certificació Europea (CE). Aquesta es una declaració jurídica vinculant del fabricant on s'exposa que el producte compleix tots els requisits del Reglament (EU) 2017/745 sobre productes sanitaris, aplicat des del maig de 2020 i que revoca la Directiva 93/42 [128].

A més de la CE, hi ha una classificació dels dispositius mèdics a través de les classificacions de la directiva europea 93/42/CEE. Segons aquesta, existeixen quatre categories en funció del seu risc potencial per la salut:

- Classe I: amb el risc més baix.
- Classe II: amb el risc potencial moderat/elevat.
- Classe III: amb el risc elevat/importat.
- Classe IV: amb el risc potencial més elevat.

1.8 Validació clínica d'algorismes d'intel·ligència artificial

La validació és un procés que es realitza un cop l'algorisme ha estat entrenat, on s'analitza la seva capacitat amb un conjunt de dades de prova, és a dir, es busca confirmar que l'algorisme

és eficaç i compleix amb l'objectiu fixat [129]. Per poder desenvolupar des de l'inici un algorisme d'IA cal seguir, resumidament, els següents passos:

- Definir a què ha de donar resposta l'algorisme o quin és el suport que es necessita que aportï.
- Fer una cerca per considerar si ja hi ha altres aplicacions al mercat comercialitzades.
- Construir l'algorisme obtenint grans conjunts de dades per entrenar-lo i per validar-lo.
- Compartir tota la informació d'aquest amb qui el farà servir per poder donar la màxima transparència.

Tots els passos del desenvolupament dels algorismes d'IA són igual de costosos i importants, però actualment, s'està posant un especial interès en el procés de validació donada la seva repercussió clínica. Al desenvolupar un algorisme d'IA, aquest necessita ser validat clínicament per confiar en la seva utilitat, per conèixer la seva capacitat, per detectar les àrees on es necessita millores i per poder comparar la capacitat amb altres models.

Existeixen dues validacions clíniques diferents i necessàries un cop l'algorisme ha estat entrenat, en funció de a quin punt es troba l'algorisme o producte desenvolupat; la validació interna i la validació externa [130]. Si el conjunt de dades per fer la validació és extret del mateix conjunt que s'ha fet servir per l'entrenament, habitualment retrospectiu, aleshores es tracta d'una validació interna. En canvi, si el conjunt de dades utilitzat és independent al conjunt de dades d'entrenament, i habitualment prospectiu, aleshores parlem de validació externa.

És important realitzar les dues validacions en el seu moment per evitar un sobreajust del model i per tant, una inflació en l'estimació de les mesures d'avaluació de la validació. Quan la validació es realitza amb un conjunt de dades conegudes, validació interna, els resultats solen ser bons però no sempre es mantenen quan la validació es realitza amb dades desconegudes, validació externa [113,130]. A més, per poder detectar amb certesa els àmbits amb major i menor eficàcia a la pràctica clínica real, cal assegurar que les dades amb les que

es realitza la validació contenen informació representativa de tota la població diana a la que s'aplicarà l'algorisme, representant totes les característiques possibles de la població.

Tot i ser conceptes clars, és necessari estipular una definició objectiva i clara sobre el terme de validació, perquè hi ha incoherències en la terminologia que provoquen poca transparència i confiança en alguns estudis. Diversos autors utilitzen el terme de validació per la prova final del model, mentre que d'altres l'utilitzen per referir-se a l'ajust del model durant el seu desenvolupament. Això pot crear confusió ja que no deixa clar si el conjunt de dades de validació ha sigut independent o no a les dades d'entrenament i per tant, no es pot saber si es tracta d'una validació interna o externa [113].

De fet, alguns estudis han suggerit establir una nomenclatura uniforme i homogènia pels conjunts de dades i tipus de validació dels algorismes [113,131]. Segons la base de dades utilitzada:

- Conjunt de dades d'entrenament, per aquelles dades utilitzades per l'entrenament.
- Conjunt de dades d'ajust, per aquelles dades utilitzades per l'ajust.
- Conjunt de dades de prova i validació, per aquelles utilitzades per l'avaluació del rendiment de l'algorisme.

Segons el conjunt de dades utilitzat per la validació:

- Validació interna, per les validacions dins de la mostra.
- Validació temporal, per les validacions dins de la mostra però amb divisió temporal.
- Validació externa, per validacions amb dades independents a la mostra.

A la literatura s'ha demostrat que la majoria d'algorismes i aplicacions d'IA utilitzades a la pràctica clínica real presenten una menor eficàcia i capacitat que els resultats obtinguts de validacions internes [116,117,132–134]. De fet, actualment la FDA de EE.UU ja exigeix que les validacions d'aplicacions d'IA es desenvolupin amb dades prospectives [130].

3. HIPÒTESIS I OBJECTIUS

Per tot això, s'exposen a continuació les hipòtesis i els objectius de la present tesi doctoral considerant la importància de la preparació, l'impacte de la implementació i les necessitats de la població i dels professionals sanitaris respecte la IA per tal de poder desenvolupar estratègies de formació i millora de l'acceptació i confiança, i recordar la necessitat d'elaborar validacions externes dels algorismes entrenats amb IA per poder assegurar el seu bon funcionament amb dades i condicions del món real.

3.1 Hipòtesis

1. Existeix una desconfiança en vers l'ús de la intel·ligència artificial en els procediments mèdics per part dels professionals d'atenció primària, provocada pel desconeixement sobre aquesta.
2. Existeix una desconfiança en vers l'ús de la intel·ligència artificial en els procediments mèdics, en especial en l'àmbit de la radiologia, per part dels usuaris d'atenció primària provocada pel desconeixement sobre aquesta.
3. Una aplicació d'intel·ligència artificial per a la lectura de radiografies de tòrax és capaç de realitzar la interpretació de la imatge tan acuradament com el radiòleg.

3.2 Objectius

3.2.1 Objectius generals

1. Conèixer la percepció i coneixement que tenen els professionals sanitaris d'atenció primària sobre l'ús de la intel·ligència artificial com a eina de suport sanitària i el seu impacte en la radiologia a través d'un qüestionari validat.
2. Conèixer la percepció i coneixement de la població sobre l'ús de la intel·ligència artificial com a eina de suport sanitària en l'àmbit de la radiologia a través d'un qüestionari validat.

3. Realitzar una validació externa amb dades prospectives d'un algorisme d'intel·ligència artificial per la lectura de radiografies de tòrax com a eina d'ajuda en el diagnòstic en l'àmbit d'atenció primària.

3.2.2. Objectius específics

1. Analitzar la percepció i el coneixement, tant dels usuaris com dels professionals sanitaris, respecte l'ús de la intel·ligència artificial en funció de les principals variables sociodemogràfiques de la mostra per identificar possibles factors relacionats.
2. Determinar quines son les principals fortaleces i preocupacions respecte l'ús de la intel·ligència artificial per poder proposar noves estratègies per tal d'augmentar el coneixement i la confiança, tant dels usuaris com dels professionals sanitaris.
3. Validar l'algorisme d'intel·ligència artificial per casos amb una única patologia i per casos amb múltiples patologies.
4. Descriure possibles millores de l'algorisme en la terminologia utilitzada i detectar possibles patologies no contemplades per l'algorisme però identificades en el món real segons el radiòleg de referència.

4. MATERIAL I MÈTODES

Pel desenvolupament de la present tesi doctoral s'ha realitzat dos estudis transversals a través de dues enquestes validades i un estudi prospectiu. Tot i que la metodologia utilitzada es troba descrita als articles publicats i adjuntats més endavant, a continuació se'n detallen alguns dels punts més importants.

4.1 Enquesta validada a la ciutadania

Amb l'objectiu de poder obtenir resultats vàlids i fiables, es va fer una cerca d'enquestes validades que avaluessin la percepció i coneixements de la ciutadania sobre la IA i el seu ús en radiologia. De la cerca, es va seleccionar l'enquesta de YP Ongena [135].

Aquesta enquesta es va desenvolupar per un equip d'investigadors d'un hospital dels Països Baixos l'any 2019. Per desenvolupar l'enquesta es va realitzar entrevistes semi estructurades a 20 participants a partir de les quals es va identificar les dimensions que van definir el marc de l'enquesta. Un cop identificades, els investigadors van desenvolupar un mínim de 7 preguntes d'escala Likert per cada dimensió inspirades en una enquesta validada sobre la orientació al canvi [136].

Amb la primera versió de l'enquesta, es va realitzar una prova pilot amb 21 pacients d'edat entre 35 i 76 anys, la meitat homes. Es va demanar als participants que responguessin l'enquesta pensant en veu alta perquè els investigadors poguessin detectar dubtes i incerteses. Amb aquesta prova pilot, es van fer canvis ajustant la terminologia de les preguntes per fer-les més entenedores i finalment, amb l'anàlisi factorial exploratori es van determinar cinc dimensions i es van obtenir valors de l'Alfa de Cronbach superiors a 0.57 en totes elles.

Un cop seleccionada l'enquesta, es va contactar amb l'autor de correspondència de l'article per tal de demanar-li l'enquesta original, l'acceptació de la traducció al català i posteriorment

el seu ús. Un cop es va obtenir la resposta positiva de l'autor i es va disposar de l'enquesta completa i original, dos investigadors van traduir-la de forma independent, la van posar en comú i un tercer va ajudar a arribar al consens en els casos de discrepàncies en la traducció. Es pot consultar l'enquesta original i l'enquesta traduïda al català als Annexos I i II.

Tot i que l'enquesta original conté 5 dimensions, l'enquesta publicada pels autors i per tant, la utilitzada, només n'aborda 4: desconfiança i responsabilitat (15 preguntes), interacció personal (6 preguntes), eficiència (4 preguntes) i estar informat (4 preguntes). Totes les preguntes son de tipus Likert (1: totalment en desacord, 5 totalment d'acord). A més, conté 5 preguntes descriptives, de tipus Likert (1: totalment en desacord, 5 totalment d'acord), sobre l'ús dels ordinadors com a eina en l'àmbit de la salut. Addicionalment, vàrem afegir una primera part sociodemogràfica a l'enquesta (sexe, edat, estat civil, nivell d'estudis i codi postal de la població de residència) i dues preguntes sobre el coneixement de la IA. Aquestes dues últimes preguntes es van afegir per donar resposta a uns dels objectius de l'estudi que era descriure el coneixement sobre la IA de la població.

Per tal de fer arribar l'enquesta a la ciutadania es van utilitzar tres vies. La primera via va ser a través dels Centres d'Atenció Primària de la Catalunya Central (compresa per les comarques de l'Anoia, el Bages, el Berguedà, Osona i el Moianès) amb una enquesta en format paper que es podia trobar als taulells dels centres o ve a través d'un codi QR que es podia trobar als cartells penjats a les sales d'espera dels centres. La segona via va ser a través de l'hospital de referència de la Catalunya Central, Hospital Althaia, on es va utilitzar, igual que als Centres d'Atenció Primària, l'enquesta en format paper i els cartells amb el codi QR. La tercera i última via va ser a través del boca a boca, compartint el codi QR a través de xarxes socials i en esdeveniments científics. Es pot consultar el cartell amb el codi QR a l'Annex III.

4.2 Enquesta validada als professionals

Després de fer una cerca de les actuals enquestes validades per avaluar la percepció i coneixement dels professionals sanitaris d'infermeria i medicina respecte la IA, es va seleccionar l'enquesta SHAIP (*Shinners Artificial Intelligence Perception*) [137].

Aquesta enquesta la va desenvolupar un equip d'investigadors d'Austràlia, l'any 2021, mitjançant el mètode *Delphi* per elaborar i validar qüestionaris. El mètode *Delphi* permet obtenir un consens sobre l'opinió d'un grup d'experts a través de debats estructurats passant així de la diversitat del grup a una única opinió representativa. Concretament, l'estudi va realitzar tres rondes a un grup de professionals sanitaris i professionals del sector informàtic en àmbit sanitari i va fixar un acord del grup superior al 75% per acceptar cada pregunta proposada.

El criteri de selecció dels professionals sanitaris per formar part del panell d'experts va ser que fossin professionals sanitaris registrats a l'Agència Australiana de Regulació de Professionals Sanitaris, amb més de deu anys d'experiència clínica en aguts o atenció primària, procedents de diferents disciplines i amb interès en la tecnologia sanitària. El criteri de selecció dels professionals del sector informàtic per formar part del panell d'experts va ser que tinguessin més de deu anys d'experiència en projectes de l'àmbit sanitari.

La primera ronda va ser a través d'una enquesta online amb l'objectiu d'identificar els temes i l'estructura. L'enquesta proposada estava formada per 18 preguntes basades en el qüestionari validat *Robot Use Self-Efficacy in Healthcare work* (RUSH) i, a més, disposava de camp lliure per afegir comentaris [138]. La segona ronda va ser a través de reunions online amb grups reduïts on es va discutir sobre el primer esborrany de l'enquesta. La tercera i última ronda va ser a través de correu electrònic per revisar i fer comentaris finals de l'enquesta.

Un cop seleccionada l'enquesta SHAIP, es va contactar amb l'autor de correspondència de l'article per tal de demanar-li l'enquesta original, l'acceptació de la traducció al català i posteriorment del seu ús. Un cop es va obtenir la resposta positiva de l'autor i es va disposar de l'enquesta completa i original, dos investigadors van traduir-la de forma independent, la

van posar en comú i un tercer va ajudar a arribar al consens en els casos de discrepàncies en la traducció. Es pot consultar l'enquesta traduïda al català i l'enquesta original als Annexos IV i V.

L'enquesta original conté deu preguntes tipus Likert que s'agrupen per avaluar dues dimensions definides, una sobre l'impacte de la IA als professionals i l'altra sobre la preparació dels professionals per l'aplicació de la IA. A més a més, es van incloure dues preguntes tancades sobre els reptes de la incorporació de la IA i la formació que els agradaria rebre, i es va afegir una primera part per obtenir informació sociodemogràfica (sexe, edat, categoria professional i anys d'experiència), així com preguntes sobre el coneixement del concepte de IA, la formació i o preparació sobre aquesta i el seu impacte en la radiologia.

Per tal de fer arribar l'enquesta als professionals es va utilitzar el correu corporatiu i es va fer un enviament massiu a tots als 1.068 professionals de medicina i infermeria de la Catalunya Central. En el correu hi havia l'enllaç per poder respondre l'enquesta en línia. Passades dues setmanes del primer enviament, es van realitzar dos enviaments posteriors com a recordatori per poder arribar al màxim de respostes possibles.

4.3 Estudi prospectiu per la validació de l'algorisme

El circuit d'interpretació de les radiografies de tòrax varia en funció del territori i del centre on es realitza la imatge. En àmbits hospitalaris hi ha un equip de tècnics de radiologia per fer les radiografies i un equip de radiòlegs destinats, entre d'altres tasques, a interpretar les imatges *in situ* i amb un curt interval de temps. En canvi, en l'àmbit de l'atenció primària només es fan radiografies en els grans centres i el sistema d'interpretació d'aquestes segueix uns circuits establerts en funció del territori.

En particular, a la Catalunya Central, compresa per les comarques de l'Anoia, el Bages, el Berguedà, el Moianès i Osona, hi ha tres centres amb capacitat per realitzar imatges radiològiques. Aquests centres tenen els anomenats Serveis d'Imatge Pel Diagnòstic (SDPI) i es troben ubicats dins un dels Centres d'Atenció Primària (CAP) de Manresa per donar servei

als usuaris del Bages, Berguedà i Moianès, d'Igualada per donar servei als usuaris d'Anoia, i de Vic per donar servei als usuaris d'Osona.

Tot i això, i respecte les radiografies de tòrax, fins l'any 2022 només dos dels tres centres disposaven d'un radiòleg per interpretar-les. Actualment, no obstant, ja només un dels tres centres de referència disposa d'un radiòleg. Això significa que, a la major part del territori de la Catalunya Central, són els especialistes de medicina de família i comunitària els qui han d'interpretar les radiografies de tòrax realitzades i, en cas de dubte, poden realitzar una interconsulta amb un especialista de radiologia de l'hospital de referència.

La interconsulta és una eina de comunicació asíncrona entre professionals. En aquest cas, el metge de família contacta, per via telemàtica, amb el radiòleg de referència de l'hospital per exposar un dubte concret. Un cop enviada la consulta, el radiòleg dona resposta amb el mínim de temps possible. Tant la pregunta com la resposta queden integrades a la història clínica de l'usuari [139].

Per poder desenvolupar l'estudi era necessari tenir l'informe d'un radiòleg de totes les radiografies de tòrax analitzades per poder tenir un *gold standard*. Així doncs, amb la necessitat pròpia de l'estudi, es va decidir desenvolupar l'estudi al centre SDPI d'Osona per ser l'únic centre de la Catalunya Central amb interpretació de totes les radiografies de tòrax de manera remota per un equip de radiòlegs de l'Hospital Can Ruti. Aquest centre dona servei a 125.000 persones.

El primer contacte entre el servei SDPI d'Osona i l'equip investigador del projecte va ser el novembre de 2021 i un cop acceptada la participació i aprovat el comitè d'ètica es va iniciar el reclutament de pacients fins el juny del 2022. La població de referència de l'estudi va ser tota la població de la comarca d'Osona que es va realitzar una radiografia de tòrax al CAP Osona durant el període d'estudi i va acceptar participar a l'estudi. Es va incloure únicament les radiografies anteroposteriors de tòrax de persones majors de 18 anys i es va excloure les dones embarassades per no esbiaixar les interpretacions de l'algorisme d'IA.

Un cop es va acabar el reclutament, des del Servei Tècnic de l'Institut Català de la Salut de la Catalunya Central es van extreure les imatges anònimes i de manera automatitzada dels pacients i els seus corresponents informes del radiòleg no anonimitzats. Seguidament, l'equip investigador va introduir totes les imatges a l'algorisme d'IA per extreure'n la interpretació (el o els diagnòstics o la possibilitat de no haver-hi alteracions patològiques).

L'informe del radiòleg no contenia un llistat uniforme i homogeni amb les alteracions de les imatges sinó que contenia un text lliure amb la descripció del que s'observava. Per aquest motiu, tres professionals de medicina de família i comunitària van interpretar els informes dels radiòlegs de referència, sense veure les imatges per no estar suggestionats. A través de la interpretació van identificar els diagnòstics del text lliure. D'aquesta manera, es va poder tenir un llistat de diagnòstics dels radiòlegs de referència i un llistat de diagnòstics de l'algorisme d'IA.

Un cop es van tenir les interpretacions de les imatges de l'algorisme d'IA i del *gold standard*, els tres professionals de medicina de família van agrupar totes les patologies detectables a través de l'algorisme en 9 categories segons l'anatomia del tòrax per poder fer un estudi individual i agrupat. Les categories van ser: implants externs, troballes o patologia del mediastí, patologia cardíaca i vàlvules, patologia de vasos, patologia òssia, patologia pleural o de l'espai pleural, troballes o patologia abdomen superior, troballes o patologia del parènquima pulmonar i altres. El llistat complet dels diagnòstics que l'algorisme d'IA pot detectar es pot consultar a l'annex VI així com l'agrupació realitzada.

5. Articles de la tesi doctoral

5.1 Article 1: Developing an Artificial Intelligence Model for Reading Chest X-rays: Protocol for a Prospective Validation Study

Protocol

Developing an Artificial Intelligence Model for Reading Chest X-rays: Protocol for a Prospective Validation Study

Queralt Miró Catalina^{1,2}, MSc; Aina Fuster-Casanovas^{1,2}, MSc; Jordi Solé-Casals^{3,4}, PhD; Josep Vidal-Alaball^{1,2,5}, MPH, MD, PhD

¹Unitat de Suport a la Recerca de la Catalunya Central, Fundació Institut Universitari per a la Recerca a l'Atenció Primària de Salut Jordi Gol i Gurina, Sant Fruitós de Bages, Spain

²Health Promotion in Rural Areas Research Group, Gerència Territorial de la Catalunya Central, Institut Català de la Salut, Sant Fruitós de Bages, Spain

³Data and Signal Processing group, Faculty of Science, Technology and Engineering, University of Vic-Central University of Catalonia, Vic, Spain

⁴Department of Psychiatry, University of Cambridge, Cambridge, United Kingdom

⁵Faculty of Medicine, University of Vic-Central University of Catalonia, Vic, Spain

Corresponding Author:

Queralt Miró Catalina, MSc

Unitat de Suport a la Recerca de la Catalunya Central

Fundació Institut Universitari per a la Recerca a l'Atenció Primària de Salut Jordi Gol i Gurina

C/ Pica d'Estats 13-15

Sant Fruitós de Bages, 08272

Spain

Phone: 34 634810263

Email: qmiro.cc.ics@gencat.cat

Abstract

Background: Chest x-rays are the most commonly used type of x-rays today, accounting for up to 26% of all radiographic tests performed. However, chest radiography is a complex imaging modality to interpret. Several studies have reported discrepancies in chest x-ray interpretations among emergency physicians and radiologists. It is of vital importance to be able to offer a fast and reliable diagnosis for this kind of x-ray, using artificial intelligence (AI) to support the clinician. Oxipit has developed an AI algorithm for reading chest x-rays, available through a web platform called ChestEye. This platform is an automatic computer-aided diagnosis system where a reading of the inserted chest x-ray is performed, and an automatic report is returned with a capacity to detect 75 pathologies, covering 90% of diagnoses.

Objective: The overall objective of the study is to perform validation with prospective data of the ChestEye algorithm as a diagnostic aid. We wish to validate the algorithm for a single pathology and multiple pathologies by evaluating the accuracy, sensitivity, and specificity of the algorithm.

Methods: A prospective validation study will be carried out to compare the diagnosis of the reference radiologists for the users attending the primary care center in the Osona region (Spain), with the diagnosis of the ChestEye AI algorithm. Anonymized chest x-ray images will be acquired and fed into the AI algorithm interface, which will return an automatic report. A radiologist will evaluate the same chest x-ray, and both assessments will be compared to calculate the precision, sensitivity, specificity, and accuracy of the AI algorithm. Results will be represented globally and individually for each pathology using a confusion matrix and the One-vs-All methodology.

Results: Patient recruitment was conducted from February 7, 2022, and it is expected that data can be obtained in 5 to 6 months. In June 2022, more than 450 x-rays have been collected, so it is expected that 600 samples will be gathered in July 2022. We hope to obtain sufficient evidence to demonstrate that the use of AI in the reading of chest x-rays can be a good tool for diagnostic support. However, there is a decreasing number of radiology professionals and, therefore, it is necessary to develop and validate tools to support professionals who have to interpret these tests.

Conclusions: If the results of the validation of the model are satisfactory, it could be implemented as a support tool and allow an increase in the accuracy and speed of diagnosis, patient safety, and agility in the primary care system, while reducing the cost of unnecessary tests.

International Registered Report Identifier (IRRID): PRR1-10.2196/39536

KEYWORDS

artificial intelligence; machine learning; chest x-ray; radiology; validation

Introduction

Chest x-rays are currently the most commonly used type of x-rays, accounting for up to 26% of all radiographic tests performed [1-3]. This technique makes it possible to identify cardiopulmonary conditions, verify the correct positioning of devices such as pacemakers, gastric and thoracic tubes, or detect obstructed blood vessels, among others [4,5].

However, chest radiography is a complex imaging modality to interpret [6]. In fact, several studies have reported discrepancies in chest x-ray interpretations among emergency physicians and radiologists [7,8]. Therefore, it is of vital importance to be able to offer a fast and reliable diagnosis for this kind of x-ray, using artificial intelligence (AI) to support the clinician.

Radiology is one of the areas in which AI has had the greatest impact. Radiologists are medical professionals who use imaging technology to diagnose pathologies. Major advances in AI have enabled these professionals to make use of this tool to improve workflows and accuracy, thus reducing economic costs by avoiding unnecessary tests [5,9].

AI is a branch of computer science that aims to simulate tasks related to human intelligence, including processes such as learning and improvement through feedback or reasoning, using machines [10]. It is a tool capable of learning and analyzing large amounts of information, in different formats and at high speed, to aid in the accuracy and speed of diagnosis, facilitate and streamline clinical care, and support public health interventions, among many other applications [11,12]. The rapid growth of computer science and big data indicates that it is here to stay and will significantly change the practice of medicine [13].

The development of a computer system capable of interpreting thoracic x-rays as efficiently as a radiologist could be of great benefit in the clinical setting. The results of Rajpurkar et al's [14] study on the application of deep learning for chest x-ray diagnosis presents an algorithm (CheXNeXt), which performs comparably with professionals in detecting multiple thoracic pathologies.

Wu et al [2] compares the interpretations of 5 radiology residents with those of an AI algorithm and corroborates that these well-trained techniques can achieve performance levels similar to professionals. Furthermore, Ciceró et al [15] demonstrates that convolutional neural networks can be trained with data sets to classify chest x-rays and obtain clinically useful performance in the detection and exclusion of common pathologies.

Oxipit is one of the leading companies in medical image reading using AI, whose goal is to introduce advances in deep learning techniques into daily clinical practice [16]. The company has developed an AI algorithm for reading chest x-rays, available through a web platform called ChestEye.

This platform is an automatic computer-aided diagnosis system where the inserted chest x-ray is read and an automatic report is returned with a capacity to detect 75 pathologies, covering 90% of diagnoses. Thus, ChestEye allows radiologists to analyze only the most relevant x-rays [17,18].

Therefore, the main objective of the study is to perform a prospective validation of the ChestEye AI algorithm as a diagnostic decision support tool for the diagnosis of chest x-rays and to try to improve or optimize it if possible.

Methods

Design

A prospective study will be conducted to validate the AI algorithm, comparing the ChestEye AI diagnoses with the radiologists' diagnoses, which is considered the gold standard. The process will include the following steps:

1. The patient will arrive at the primary care center for the chest x-ray, and if he/she meets the inclusion and exclusion criteria, the health care staff will briefly explain the study and provide the informed consent form to be signed.
2. Regardless of whether the user has agreed to participate in the study or not, the reference radiologist will perform the diagnosis of the x-ray to be entered into the Primary Care Clinical Station (ECAP). This station is the computerized clinical history program used by all professionals in the primary care network of the Institut Català de la Salut (ICS).
3. If the user has agreed to participate in the study, the researchers will extract the ECAP x-ray and enter it into the AI algorithm through their web-based platform to obtain their diagnosis.
4. Finally, the performance and fit of the AI model against the gold standard (radiologists' diagnoses) will be validated and evaluated.

The AI algorithm ChestEye, from Oxipit, is an automatic and autonomous algorithm, without the involvement of the radiologist, which works through a web-based platform where the image is entered in DICOM format, and returns an image evaluation and diagnosis. The algorithm has the capacity to detect 75 pathologies, covering 90% of the diagnoses [16].

ChestEye has been previously developed and trained by Oxipit through iterative processing of large amounts of data by neural network-based AI algorithms, allowing the software to learn automatically from patterns or features in the data.

Scope, Period, and Participants

The study will be performed at the ICS Primary Care Centre Vic Nord (Osona, Catalonia, Spain), a reference center where all chest x-rays in the region are performed. It is expected that data can be obtained in 5 to 6 months, from February 7, 2022, with recruitment using consecutive sampling. In June 2022,

more than 450 chest x-rays have been collected, so it is expected that 600 samples will be gathered in July 2022.

The reference population of the prospective study will be the entire population of Osona due to undergo a chest x-ray at this center, with prior informed consent.

The study will include only anteroposterior chest x-rays performed from the beginning of the study until the necessary sample is obtained from patients with authorized informed consent and who are older than 18 years. Pregnant women and chest x-rays of inadequate quality (poor exposure, images not centered or rotated) will be excluded from the study as the AI algorithm needs high-quality images to maximize its performance.

Sample Size and Sampling Procedure

To validate the AI algorithm, a total sample of 600 x-rays will be needed, 200 of them with one of the 75 pathologies detected by the AI algorithm. The proposed sample is based on calculations used in similar research [1,14,19,20]. Furthermore, it has been calculated that with this sample size, we can estimate global accuracy considered to be around 70% with 95% confidence, 4% precision, and an anticipated replacement rate of 15%.

Data Collection and Information Sources

The ICS health care personnel performing the chest x-rays will explain the study and its objectives to the users, and will give the patient an information sheet, together with the informed consent form, to all those who meet the inclusion criteria. The ICS Central Catalonia technical service will then extract all these x-rays with their corresponding diagnosis. Each x-ray will be associated with a unique identifier to relate it to its diagnosis and eliminate any nonanonymized information. Next, the study's principal researchers will input the x-rays into the AI system to obtain the diagnoses of the models using the algorithm. Finally, the data will be analyzed by comparing the diagnoses of the practitioner and the algorithm.

Data Analysis

To validate the algorithm, the results using the AI algorithm and the diagnoses made by radiologists will be compared. With this, the confusion matrix of the algorithm will be obtained from the correctly classified positive (TP), correctly classified negative (TN), false positive (FP), and false negative (FN) x-rays. The sensitivity, specificity, classification rate (accuracy), and area under the curve (AUC) of the algorithm will be calculated from this matrix. These results can be obtained for each pathology and the classifier as a whole. Accuracy, recall, and F-measurement will also be calculated for the overall classifier and each pathology.

To evaluate the classifier for multipathology radiology, the data will be treated as a set of binary variables, one for each pathology. In this case, the AUC will be calculated using the One-vs-All method. Macroaveraging and microaveraging measures will be considered to highlight pathologies with lower prevalence. The data will be analyzed with the statistical software R (version 4.1.2; R Foundation for Statistical

Computing), whose intervals will be of 95% confidence, with a significance level of 5%.

Ethics Approval

The University Institute for Research in Primary Health Care Jordi Gol i Gurina (Barcelona, Spain) ethics committee approved the trial study protocol (approval code: 21/288). Written informed consent will be requested from all patients participating in the study.

Results

Patient recruitment began in February 2022, and it is expected that data can be obtained in 5 to 6 months. On June 2022, more than 450 chest x-rays have been collected, so it is expected that 600 samples will be gathered in July 2022. Each user who agrees to participate in the study will be asked for written informed consent and will be given the project information sheet. Data collection for all participants is expected to be completed by June 2022, and the results can be published by the end of 2022.

In this way, we hope to obtain sufficient evidence to demonstrate that the use of AI in the reading of chest x-rays can be a good tool for diagnostic support. However, in the context of Central Catalonia (the Catalan region where the data was collected), there is an increasingly lower volume of radiologists, and therefore, tools need to be developed to support professionals who have to interpret these tests [21,22].

Once the algorithm has been validated, the values of sensitivity, specificity, accuracy, and AUC will be used to evaluate the results obtained and to determine whether it would be a good model to be introduced in the Catalan health system.

Discussion

Comparison With Prior Work

The protocol of this study aims to perform a prospective validation of an AI algorithm and to demonstrate that the use of AI in chest x-rays can become a good tool for supporting professionals in their diagnoses. In this context, this study may bring added value for both patients and primary care physicians as it will provide information about the effectiveness of the AI algorithm and its limitations. External validation of new AI tools is essential before implementing them as diagnostic systems.

Studies are showing that the application of AI models can be comparable to the performance of a professional in the detection of multiple pathologies [2,14,15]. However, before committing resources to AI applications in health care, the acceptance of these applications should be studied. Although some studies have shown that AI has a high potential to be useful as a diagnostic tool, it is remarkable that most patients still preferred the diagnoses done by physicians, and professionals only accepted AI models if they were used in combination with "human diagnosis" [23,24]. In this context, leading health care systems are moving toward the digitization of health care. Therefore, it is time to provide and validate tools that can enable improvement in the workflow of professionals as well as support

their diagnosis. Always consider the clinical context for the subsequent application of these tools.

Furthermore, it has to be taken into consideration that most of the AI studies conducted in health care were just proof-of-concept projects that used retrospective clinical data sets [25]. The application of AI techniques in the real clinical context is becoming more and more relevant to ensure its safe adoption in health care systems. Thus, this study will be conducted using prospective data sets, promoting the health care AI researchers' community to work closely with health care providers in a real clinical environment.

Limitations

This study has some limitations. The most relevant one is that there is the possibility of not obtaining a homogeneous

distribution across the 75 possible diagnoses due to their low prevalence. In that sense, as a large number of diseases can be detected by chest x-ray, we will probably not obtain representative results for the less prevalent diseases. As class imbalance may be a limitation, the F score will be evaluated. Otherwise, the large number of more frequent pathologies may overestimate the quality of the algorithm (accuracy, sensitivity, and specificity). Another possible limitation is that a small amount of sample is likely to be lost due to inadequate image quality, as chest x-rays of inadequate quality will be excluded.

Conclusions

If the results of the model validation are satisfactory, the model can be implemented as a support tool and can increase diagnostic accuracy and speed, patient safety and agility within the primary care system, and reduce unnecessary testing costs.

Data Availability

Our manuscript is based on confidential and sensitive health data. However, to support scientific transparency, we will publish deidentified data for reviewers or for replication purposes. The data will be deposited and made available in our publicly accessible Mendeley repository.

Conflicts of Interest

None declared.

References

1. Hwang EJ, Park S, Jin K, Kim JI, Choi SY, Lee JH, et al. DLAD DevelopmentEvaluation Group. Development and validation of a deep learning-based automated detection algorithm for major thoracic diseases on chest radiographs. *JAMA Netw Open*. Mar 01, 2019;2(3):e191095. [FREE Full text] [doi: [10.1001/jamanetworkopen.2019.1095](https://doi.org/10.1001/jamanetworkopen.2019.1095)] [Medline: [30901052](https://pubmed.ncbi.nlm.nih.gov/30901052/)]
2. Wu J, Wong K, Gur Y, Ansari N, Karargyris A, Sharma A, et al. Comparison of chest radiograph interpretations by artificial intelligence algorithm vs radiology residents. *JAMA Netw Open*. Oct 01, 2020;3(10):e2022779. [FREE Full text] [doi: [10.1001/jamanetworkopen.2020.22779](https://doi.org/10.1001/jamanetworkopen.2020.22779)] [Medline: [33034642](https://pubmed.ncbi.nlm.nih.gov/33034642/)]
3. Santos ÁM, Solís PV. Posición SERAM sobre la necesidad de informar la radiología simple. *SERAM*. 2015. URL: <https://drive.google.com/file/d/1x1xnri4LbC60yxJgBx35bhEt51M-BxoT/view> [accessed 2022-10-18]
4. Radiografía: Estudio de diagnóstico por imágenes que ayuda a diagnosticar con rapidez. Mayo Clinic. URL: <https://www.mayoclinic.org/es-es/tests-procedures/x-ray/about/pac-20395303> [accessed 2022-10-18]
5. Johnson AEW, Pollard TJ, Berkowitz SJ, Greenbaum NR, Lungren MP, Deng C, et al. MIMIC-CXR, a de-identified publicly available database of chest radiographs with free-text reports. *Sci Data*. Dec 12, 2019;6(1):317. [FREE Full text] [doi: [10.1038/s41597-019-0322-0](https://doi.org/10.1038/s41597-019-0322-0)] [Medline: [31831740](https://pubmed.ncbi.nlm.nih.gov/31831740/)]
6. Gatt ME, Spectre G, Paltiel O, Hiller N, Stalnikowicz R. Chest radiographs in the emergency department: is the radiologist really necessary? *Postgrad Med J*. Apr 2003;79(930):214-217. [FREE Full text] [doi: [10.1136/pmj.79.930.214](https://doi.org/10.1136/pmj.79.930.214)] [Medline: [12743338](https://pubmed.ncbi.nlm.nih.gov/12743338/)]
7. Preston CA, Marr JJ, Amaraneni KK, Suthar BS. Reduction of "callbacks" to the ED due to discrepancies in plain radiograph interpretation. *Am J Emerg Med*. Mar 1998;16(2):160-162. [doi: [10.1016/s0735-6757\(98\)90036-5](https://doi.org/10.1016/s0735-6757(98)90036-5)] [Medline: [9517693](https://pubmed.ncbi.nlm.nih.gov/9517693/)]
8. Espinosa JA, Nolan TW. Reducing errors made by emergency physicians in interpreting radiographs: longitudinal study. *BMJ*. Mar 18, 2000;320(7237):737-740. [FREE Full text] [doi: [10.1136/bmj.320.7237.737](https://doi.org/10.1136/bmj.320.7237.737)] [Medline: [10720354](https://pubmed.ncbi.nlm.nih.gov/10720354/)]
9. Do H, Spear L, Nikpanah M, Mirmomen S, Machado L, Toscano A, et al. Augmented radiologist workflow improves report value and saves time: a potential model for implementation of artificial intelligence. *Acad Radiol*. Jan 2020;27(1):96-105. [FREE Full text] [doi: [10.1016/j.acra.2019.09.014](https://doi.org/10.1016/j.acra.2019.09.014)] [Medline: [31818390](https://pubmed.ncbi.nlm.nih.gov/31818390/)]
10. Benavent ND, Colomer J, Luis M, Gutiérrez Q, Gol-Montserrat J, del Llano Señarís JE. Inteligencia artificial y decisiones clínicas: cómo está cambiando el comportamiento del médico. *Fundación Gaspar Casal*. Dec 10, 2020. URL: <https://fundaciongasparcasal.org/inteligencia-artificial-y-decisiones-clinicas-como-esta-cambiando-el-compartimento-medico/> [accessed 2022-10-18]
11. Ramesh AN, Kambhampati C, Monson J, Drew P. Artificial intelligence in medicine. *Ann R Coll Surg Engl*. Sep 2004;86(5):334-338. [FREE Full text] [doi: [10.1308/147870804290](https://doi.org/10.1308/147870804290)] [Medline: [15333167](https://pubmed.ncbi.nlm.nih.gov/15333167/)]
12. Ávila-Tomás JF, Mayer-Pujadas MA, Quesada-Varela VJ. [Artificial intelligence and its applications in medicine II: current importance and practical applications]. *Aten Primaria*. Jan 2021;53(1):81-88. [FREE Full text] [doi: [10.1016/j.aprim.2020.04.014](https://doi.org/10.1016/j.aprim.2020.04.014)] [Medline: [32571595](https://pubmed.ncbi.nlm.nih.gov/32571595/)]

13. Karimi A, HaddadPajouh H. Artificial intelligence, important assistant of scientists and physicians. *Galen Med J*. 2020;9:e2048. [FREE Full text] [doi: [10.31661/gmj.v9i0.2048](https://doi.org/10.31661/gmj.v9i0.2048)] [Medline: [34466625](https://pubmed.ncbi.nlm.nih.gov/34466625/)]
14. Rajpurkar P, Irvin J, Ball R, Zhu K, Yang B, Mehta H, et al. Deep learning for chest radiograph diagnosis: a retrospective comparison of the CheXNeXt algorithm to practicing radiologists. *PLoS Med*. Nov 2018;15(11):e1002686. [FREE Full text] [doi: [10.1371/journal.pmed.1002686](https://doi.org/10.1371/journal.pmed.1002686)] [Medline: [30457988](https://pubmed.ncbi.nlm.nih.gov/30457988/)]
15. Cicero M, Bilbily A, Colak E, Dowdell T, Gray B, Perampaladas K, et al. Training and validating a deep convolutional neural network for computer-aided detection and classification of abnormalities on frontal chest radiographs. *Invest Radiol*. 2017;52(5):281-287. [doi: [10.1097/rli.0000000000000341](https://doi.org/10.1097/rli.0000000000000341)] [Medline: [27922974](https://pubmed.ncbi.nlm.nih.gov/27922974/)]
16. Oxipit ChestEye obtiene la certificación de dispositivos médicos en Australia. Oxipit. 2020. URL: <https://oxipit.ai/news/oxipit-ai-medical-imaging-australia/> [accessed 2022-10-18]
17. Un software con IA para las radiografías de tórax recibe la certificación CE. MedImaging. URL: <https://www.medimaging.es/industria/articulos/294777065/un-software-con-ia-para-las-radiografias-de-torax-recibe-la-certificacion-ce.html> [accessed 2022-10-18]
18. Study: AI found to reduce bias in radiology reports. Oxipit. URL: <https://oxipit.ai/news/study-ai-found-to-reduce-bias-in-radiology-reports/> [accessed 2022-10-18]
19. Nam JG, Park S, Hwang EJ, Lee JH, Jin K, Lim KY, et al. Development and validation of deep learning-based automatic detection algorithm for malignant pulmonary nodules on chest radiographs. *Radiology*. Jan 2019;290(1):218-228. [doi: [10.1148/radiol.2018180237](https://doi.org/10.1148/radiol.2018180237)] [Medline: [30251934](https://pubmed.ncbi.nlm.nih.gov/30251934/)]
20. Park S, Lee SM, Lee KH, Jung KH, Bae W, Choe J, et al. Deep learning-based detection system for multiclass lesions on chest radiographs: comparison with observer readings. *Eur Radiol*. Mar 30, 2020;30(3):1359-1368. [doi: [10.1007/s00330-019-06532-x](https://doi.org/10.1007/s00330-019-06532-x)] [Medline: [31748854](https://pubmed.ncbi.nlm.nih.gov/31748854/)]
21. Martín I. Martí de Gracia: Vivimos una situación crítica de escasez de radiólogos. Redacción Médica. URL: <https://www.redaccionmedica.com/secciones/radiologia/marti-de-gracia-vivimos-una-situacion-critica-de-escasez-de-radiologos--4663> [accessed 2022-06-25]
22. Esquerrà M, Poch PR, Ticó TM, Canal V, Mir JM, Cruxent R. [Abdominal ultrasound: a diagnostic tool within the reach of general practitioners]. *Aten Primaria*. Oct 2012;44(10):576-583. [FREE Full text] [doi: [10.1016/j.aprim.2011.07.016](https://doi.org/10.1016/j.aprim.2011.07.016)] [Medline: [22018792](https://pubmed.ncbi.nlm.nih.gov/22018792/)]
23. Lennartz S, Dratsch T, Zopfs D, Persigehl T, Maintz D, Hokamp NG, et al. Use and control of artificial intelligence in patients across the medical workflow: single-center questionnaire study of patient perspectives. *J Med Internet Res*. Feb 17, 2021;23(2):e24221. [FREE Full text] [doi: [10.2196/24221](https://doi.org/10.2196/24221)] [Medline: [33595451](https://pubmed.ncbi.nlm.nih.gov/33595451/)]
24. Maassen O, Fritsch S, Palm J, Deffge S, Kunze J, Marx G, et al. Future medical artificial intelligence application requirements and expectations of physicians in German university hospitals: web-based survey. *J Med Internet Res*. Mar 05, 2021;23(3):e26646. [FREE Full text] [doi: [10.2196/26646](https://doi.org/10.2196/26646)] [Medline: [33666563](https://pubmed.ncbi.nlm.nih.gov/33666563/)]
25. Yin J, Ngiam KY, Teo HH. Role of artificial intelligence applications in real-life clinical practice: systematic review. *J Med Internet Res*. Apr 22, 2021;23(4):e25759. [FREE Full text] [doi: [10.2196/25759](https://doi.org/10.2196/25759)] [Medline: [33885365](https://pubmed.ncbi.nlm.nih.gov/33885365/)]

Abbreviations

AI: artificial intelligence

AUC: area under the curve

ECAP: Estació Clínica d'Atenció Primària (Primary Care Clinical Station)

ICS: Institut Català de la Salut

Edited by T Leung; submitted 13.05.22; peer-reviewed by F Seguí, R Rastmanesh, Z Li; comments to author 08.06.22; revised version received 27.06.22; accepted 08.07.22; published 16.11.22

Please cite as:

Miró Catalina Q, Fuster-Casanovas A, Solé-Casals J, Vidal-Alaball J

Developing an Artificial Intelligence Model for Reading Chest X-rays: Protocol for a Prospective Validation Study

JMIR Res Protoc 2022;11(11):e39536

URL: <https://www.researchprotocols.org/2022/11/e39536>

doi: [10.2196/39536](https://doi.org/10.2196/39536)

PMID: [36383419](https://pubmed.ncbi.nlm.nih.gov/36383419/)

©Queralt Miró Catalina, Aina Fuster-Casanovas, Jordi Solé-Casals, Josep Vidal-Alaball. Originally published in JMIR Research Protocols (<https://www.researchprotocols.org>), 16.11.2022. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution,

and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on <https://www.researchprotocols.org>, as well as this copyright and license information must be included.

5.2 Article 2: Knowledge and perception of primary care healthcare professionals on the use of artificial intelligence as a healthcare tool

Knowledge and perception of primary care healthcare professionals on the use of artificial intelligence as a healthcare tool

DIGITAL HEALTH
Volume 9: 1–11
© The Author(s) 2023
Article reuse guidelines:
sagepub.com/journals-permissions
DOI: 10.1177/20552076231180511
journals.sagepub.com/home/dhj



Queralt Miró Catalina^{1,2} , Aïna Fuster-Casanovas^{1,2},
Josep Vidal-Alaball^{1,2,3}, Anna Escalé-Besa^{1,2}, Francesc X Marin-Gomez^{1,2},
Joaquim Femenia³ and Jordi Solé-Casals^{4,5}

Abstract

Objective: The rapid digitisation of healthcare data and the sheer volume being generated means that artificial intelligence (AI) is becoming a new reality in the practice of medicine. For this reason, describing the perception of primary care (PC) healthcare professionals on the use of AI as a healthcare tool and its impact in radiology is crucial to ensure its successful implementation.

Methods: Observational cross-sectional study, using the validated Shinners Artificial Intelligence Perception survey, aimed at all PC medical and nursing professionals in the health region of Central Catalonia.

Results: The survey was sent to 1068 health professionals, of whom 301 responded. And 85.7% indicated that they understood the concept of AI but there were discrepancies in the use of this tool; 65.8% indicated that they had not received any AI training and 91.4% that they would like to receive training. The mean score for the *professional impact of AI* was 3.62 points out of 5 (standard deviation (SD) = 0.72), with a higher score among practitioners who had some prior knowledge of and interest in AI. The mean score for *preparedness for AI* was 2.76 points out of 5 (SD = 0.70), with higher scores for nursing and those who use or do not know if they use AI.

Conclusions: The results of this study show that the majority of professionals understood the concept of AI, perceived its impact positively, and felt prepared for its implementation. In addition, despite being limited to a diagnostic aid, the implementation of AI in radiology was a high priority for these professionals.

Keywords

Artificial intelligence, knowledge, perception, survey, technology, digital health, primary care

Submission date: 30 March 2023; Acceptance date: 19 May 2023

Introduction

Artificial intelligence (AI) is defined as ‘the ability of machines to mimic human behaviour’.¹ It is a set of mathematical models, expressed in the form of algorithms, capable of learning and analysing large amounts of data, in different formats and at high speed, to help in the accuracy and speed of diagnosis, facilitate and streamline clinical care, and support public health interventions, among many other applications in the field of health.^{2,3} Despite being recent concepts, the rapid digitisation of healthcare data and the sheer volume being generated means that artificial intelligence (AI) is becoming a new reality in the practice of medicine.⁴

¹Unitat de Suport a la Recerca de la Catalunya Central, Fundació Institut Universitari per a la Recerca a l’Atenció Primària de Salut Jordi Gol i Gurina, Sant Fruitós de Bages, Spain

²Health Promotion in Rural Areas Research Group, Gerència Territorial de la Catalunya Central, Institut Català de la Salut, Sant Fruitós de Bages, Spain

³Faculty of Medicine, University of Vic-Central University of Catalonia, Vic, Spain

⁴Data and Signal Processing group, Faculty of Science, Technology and Engineering, University of Vic-Central University of Catalonia, Vic, Spain

⁵Department of Psychiatry, University of Cambridge, Cambridge, UK

Corresponding author:

Josep Vidal-Alaball, Unitat de Suport a la Recerca de la Catalunya Central, Fundació Institut Universitari per a la Recerca a l’Atenció Primària de Salut Jordi Gol i Gurina, Sant Fruitós de Bages, 08272, Spain.

Email: jvidal.cc.ics@gencat.cat



Currently, AI is already able to support in patient triage, suggest diagnoses and result alerts,⁵ process natural language to automate clinical documentation by text summarisation,⁶ make diagnoses from radiological images⁷ or drawings such as Archimedes' spiral,⁸ etc. Several studies have demonstrated the potential usefulness of these AI algorithms in many medical specialties such as ophthalmology,⁹ dermatology,¹⁰ neurology¹¹ and radiology,¹² the latter being one of the healthcare fields where AI has had a greater impact, especially in chest radiographs.^{13,14}

Despite being one of the most widely used typologies,^{15,16} chest radiography is a complex modality to interpret¹⁷ and often creates discrepancies between the radiologist's and family physician's assessment.^{18,19} The region of Central Catalonia has the particularity of having few radiologists. As a result, primary care (PC) physicians are increasingly required to interpret X-rays. For this reason, the use of AI algorithms capable of interpreting thoracic radiographs approaching the efficiency of a radiologist could represent a great benefit in the clinical setting as a diagnostic support tool.

Knowing how healthcare professionals perceive the use of AI in the healthcare setting is of vital importance in order to tailor the technology to the needs of professionals, align it with their values and make it easy to use.^{20–23} Several studies have detected hope and optimism among healthcare professionals in its application, but also some concern about its ethical implications, lack of knowledge and training on its use, which may contribute to some reluctance among professionals.^{24–26}

Diaz et al.,²⁷ through a survey, observed that the majority of healthcare professionals agree that AI will improve their day-to-day lives, but that work needs to be done on the need for specific AI training. An integrative review by Shinnars et al.²⁸ showed that healthcare professionals were less likely to use AI if they did not trust the technology or understand how it was used to improve patient outcomes. The study by Sit et al.²⁹ demonstrated that students were aware of the importance of AI in their careers and selected the radiology specialty less due to the perception that AI would replace them in the future. Thus, it is necessary to introduce AI directly into the academic training of healthcare professionals so that they can enter the working world with knowledge of this tool.^{30,31}

Jha et al.,³² when asked about the impact of AI in primary healthcare information systems, observed that physicians were sceptical about the ability of AI to perform better than humans. Another study³³ found that healthcare professionals share concerns about data management, knowledge development, deterioration of the doctor-patient relationship, and alteration of the diagnostic and decision-making landscape.

AI applied to healthcare has great potential, although analysing the perception of stakeholders is key to ensuring its correct implementation. Thus, the aim of this study is to

describe the perception and knowledge of PC healthcare professionals on the use of AI as a healthcare tool, and its impact in radiology, through a validated survey. The results obtained will be essential to develop strategies aimed at increasing the acceptance of AI use and reducing possible resistance to change.

Methods

Study design and sample

Observational cross-sectional study aimed at all PC medical and nursing professionals in the Central Catalonia Health Region. This region includes the counties of Anoia, Bages, Berguedà, Moianès and Osona, and has approximately 1000 professionals. The survey was sent via corporate email to all these professionals up to three times to obtain the maximum response rate. The period during which the survey was active was from 2 November 2022 to 31 January 2023. The survey was open to all PC medical and nursing professionals of the Catalan Institute of Health who were over 18 years of age and accepted the invitation they received by email. Responses from other professional categories have been excluded.

A minimum of 283 surveys, distributed in the study region were required, in order to estimate with a 95% confidence and a precision of 0.1 points, the values of the two dimensions of the questionnaire, assuming a standard deviation of 1 point.²²

Sources

A validated survey known as Shinnars Artificial Intelligence Perception (SHAIP)³⁴ was used. The survey deals with two dimensions defined in order to describe *Professionals impact of AI* and *Preparedness for AI* with 10 questions using the Likert scale (1 means strongly disagree and 5 means strongly agree). Additionally, there are two descriptive multiple-choice questions about the challenges of incorporating AI and the training they would like to receive, and an initial part was added to obtain sociodemographic information (sex, age, professional category and years of professional experience), as well as questions related to knowledge of the concept of AI, one closed and the other open, training and/or preparation for it, and its impact on radiology.

Permission was requested from the author to use it and translate it into Catalan. Two researchers translated it separately and pooled it, and a third helped to reach consensus in cases of discrepancies in translation. The survey was sent to health professionals in digital format.

Statistical analysis

Categorical variables have been described with absolute frequency and percentage, and continuous variables with

Table 1. Descriptive analysis of the sample.

	Total sample (N = 301)
How old are you, in years (N = 292)**	46.0 (11.0)*
Categorised age (N = 292)	
[18-30)	22 (7.53%)
[30-40)	62 (21.2%)
[40-50)	101 (34.6%)
[50-60)	70 (24.0%)
> = 60	37 (12.7%)
What is your gender?	
Female	244 (81.1%)
Male	57 (18.9%)
What health discipline do you work in?	
Nursing	144 (47.8%)
Medicine	146 (48.5%)
Others	11 (3.65%)
How many years of experience do you have in your job?	
Less than 1 year	11 (3.65%)
Between 1 and 2 years	24 (7.97%)
Between 3 and 4 years	43 (14.3%)
Between 5 and 10 years	35 (11.6%)
More than 10 years	188 (62.5%)
Do you understand what the concept of AI means?	
No	43 (14.3%)
Yes	258 (85.7%)
In your work do you use AI to perform or plan care?	
No	99 (32.9%)

(continued)

Table 1. Continued.

	Total sample (N = 301)
I do not know	71 (23.6%)
Yes	131 (43.5%)
AI education or training:	
None	198 (65.8%)
Self-starting online courses, webinars, conferences, etc.	63 (20.9%)
Job training	30 (9.97%)
University education	10 (3.32%)
Would you like to receive education on AI in health?	
No	26 (8.64%)
Yes	275 (91.4%)

Absolute frequency and percentage; n (%).

* Mean and standard deviation. ** Only 292 professionals answered the age correctly.

AI: artificial intelligence.

mean and standard deviation. Factor analysis (FA) and Cronbach's alpha was used to determine the validity and reliability of two dimensions. In general, Cronbach's alpha of 0.7 is taken as indication of good internal consistency. In some cases, an alpha of 0.5 or 0.6 can be acceptable.^{35,36} For the bivariate analysis between the dimensions and the sociodemographic variables, the Student's t-test or analysis of variance with multiple comparisons was used. The analyses were performed with R statistical software version 4.2.1, and the significance level was set at 5%.

Ethical considerations

The study protocol was approved by the University Institute for Primary Care Research (IDIAP) Jordi Gol Health Care Ethics Committee (Code 20/177-PCV). The survey was anonymous, so the patient consent was not necessary.

Results

The survey was sent to 1068 health professionals, of whom 301 responded, obtaining a response rate of 28.2%.

Of the professionals who responded to the survey, the mean age was 46.0 years (SD = 11.0) with 81.1% female,

47.8% were nursing professionals and 48.5% were medical professionals, the vast majority with more than 10 years of professional experience in their current job (62.5%) (Table 1).

With regard to the use and knowledge of AI, 85.7% indicated that they understood its concept but saw discrepancies in its use due to the execution and/or planning in their work; 43.5% indicated that they used it, 32.9% that they did not and 23.6% that they did not know. More than half of the respondents indicated that they had not received any training on AI (65.8%) and the vast majority responded that they would like to receive it (91.4%) (Table 1).

Figure 1 describes the main challenges of incorporating AI into the workplace and topics they would like to know more about. On the one hand, the main incorporation challenges were the knowledge and skills of the professional staff (58.5%), interoperability with current systems (48.2%), infrastructure (46.8%), implementation cost (44.2%) and organisational support (43.2%). On the other hand, the main training topics were AI application in healthcare (85.0%), AI ethics in healthcare (50.8%) and general education on AI capabilities (36.2%).

Across the SHAIIP questionnaire items (1: strongly disagree and 5: strongly agree), two dimensions were obtained; *professional impact of AI* as the mean of the individual scores of questions 1 to 6, and the *preparedness for AI* as the mean of the individual scores of questions 7 to 10. In a preliminary analysis, we examined the Cronbach's alpha estimation of internal consistency and reliability of two dimensions. The Cronbach's alpha estimation of *professional impact of AI* was 0.84 (95% confidence interval (CI) 0.79–0.88) which suggests that items had very little variance specific to individual items. The Cronbach's alpha estimation of *preparedness for AI* was a 0.56 (95% CI 0.46–0.65). Figure 2 shows the results of the FA to validate the two dimensions of survey with study data and study language.

Table 2 describes the two dimensions related to the professional impact of AI and the preparedness for AI. Professionals scored different items (on a Likert scale) related to each dimensions. The mean score of responses on the *professional impact of AI* was 3.62 points (standard deviation (SD) = 0.72) out of 5 on the Likert scale. Thus, they consider that AI will improve their workflow, population health indicators and the sustainability of the healthcare system. Specifically, the two items with the highest scores meaning that they agree were 'I believe that the use of AI in my specialty could improve clinical decision making' (3.87 out of 5 mean score with SD = 0.87) and 'I believe that AI can improve population health indicators' (3.81 out of 5 mean score with SD = 0.86). In contrast, the two lowest scoring items indicated that they were unsure were 'I believe that one day AI will be able to be part of my job as a healthcare professional' (3.40 out of 5 mean score with SD = 1.10) and 'I believe that the introduction of AI can reduce the economic cost associated with my work' (3.37 out of 5 mean score with SD = 0.97) (Table 2).

On the other hand, the mean score of the responses on the *preparedness for AI* was 2.76 points (SD = 0.70) out of 5 on the Likert scale. Thus, professionals are unsure about if they are well prepared to use IA in their workflow, yet they assume their responsibility in the use of AI. Of the four items in the dimension, the one with the highest score meaning that they agree was 'I believe there is an ethical framework for the use of AI technology in my workplace' (3.23 out of 5 mean score with SD = 1.05) and the one with the lowest score indicating that they were unsure was 'I believe I have been adequately trained to use AI in my job' (2.02 out of 5 mean score with SD = 1.01) (Table 2).

Table 3 shows the results of the bivariate analysis between the two dimensions and the main explanatory variables. With regard to the *professional impact of AI*,

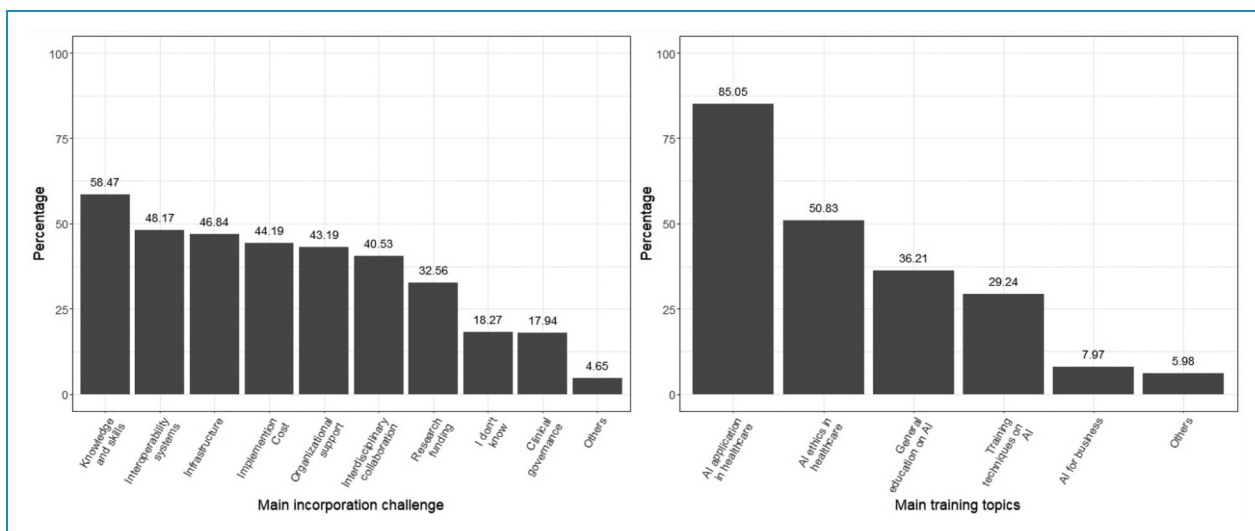


Figure 1. Challenges of incorporating artificial intelligence into the workplace and topics they would like to know more about.

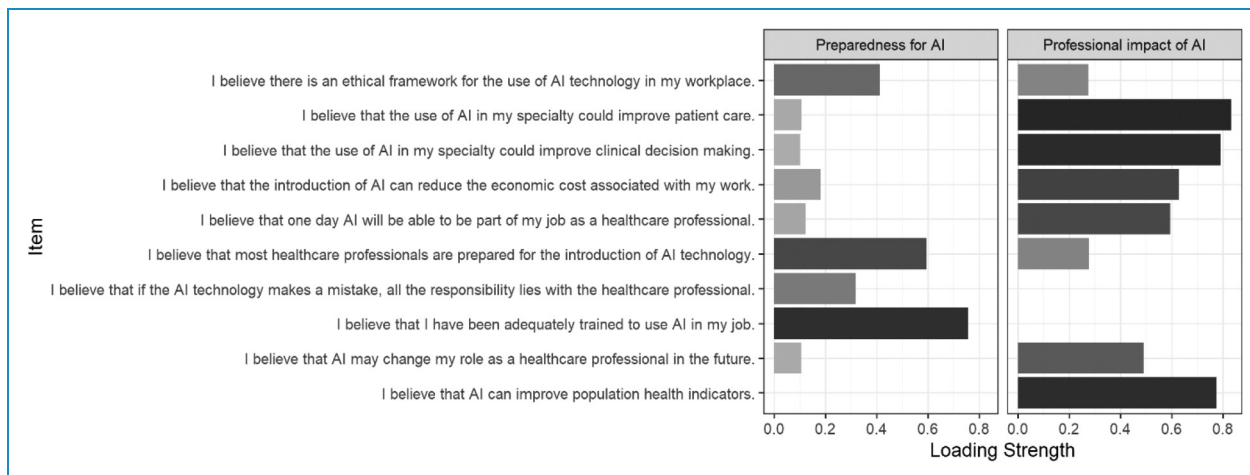


Figure 2. Factor analysis to validate dimensions of questionnaire.

healthcare professionals who understand the concept of the AI reported a greater impact of AI than those who did not. Those who have received training such as self-starting online courses, webinars, lectures or conferences have significantly different perceptions ($p = 0.012$) of the AI impact to their professional role compared with those that have not had any training. In addition, professionals who want to receive training have reported a greater impact of the AI than those who did not. Otherwise, nursing professionals' perception of preparedness for AI is significantly different than the ones from medical professionals, with nurses feeling more prepared than doctors ($p = 0.024$). Furthermore, professionals who are using AI or did not know have reported that they feel more prepared for AI than those who did not.

At the end of the survey, they were asked about the priority level for three possible applications of AI in radiology. This was because radiology is one of the specialties where AI has made a significant impact and the region of Central Catalonia is currently in the process of validating AI algorithms for use in radiology. Of the three possible applications, the high percentage of professionals who consider it a high priority to implement AI for image interpretation to obtain immediate preliminary results (41.2%) and the high percentage of professionals who consider that AI should not be implemented for image interpretation to obtain final and definitive patient results (18.9%) stands out (Figure 3).

Discussion

The aim of this study was to characterise the understanding and perception of PC healthcare professionals regarding the application of AI as a tool in healthcare, and to assess its potential impact on radiology. The findings of the present study provide insight into the perception and knowledge of AI as a healthcare tool among PC healthcare professionals in the region of Central Catalonia. Conducting such analyses is relevant for developing future strategies

aimed at improving the acceptance and application of AI in a specific region.

Regarding the professional's knowledge of the concept of AI, most of them feel that they understand what it means, and the majority indicates that they do use it in their work. However, although the Catalan Institute of Health, a health provider for 5.8 million users in Catalonia,³⁷ is working on the validation of AI models applied in imaging with the aim of implementing it in consultations^{38–40} and is in an incipient development of different applications,⁴¹ currently there are still none integrated in the patient's clinical history. Therefore, although the vast majority indicates that they are aware of the concept of AI, there is a general lack of knowledge about its applicability and its possible uses.

In fact, Chen found in the meta-analysis⁴² that only 10% to 30% of all healthcare professionals used AI in their clinical practice. Moreover, these results are consistent with the results of the meta-analysis⁴² where it was observed that 74.29% of the included studies indicated a lack of knowledge of the professionals and 15 studies suggested the urgent need to integrate AI into the training programmes of residents. Also, a survey of 3018 medical students in Turkey showed that only 2.8% felt well informed about the use of AI in medicine,⁴³ and a survey answered by 484 UK medical students showed that 89% of students felt that receiving AI training would be beneficial to their career, as well as 78% believing that AI training was necessary as part of a university degree.²⁹ Otherwise, various studies in the global literature suggest that incorporating AI education is crucial in overcoming barriers and limitations in the use and understanding of AI.^{44–47}

In terms of professionals' perceptions, a positive impact of AI was reported in the survey (scored 3.62 out of 5) similar to that obtained in the Shinnars et al.'s²² study (scored 3.66 out of 5). This agreement was higher for professionals who had some prior knowledge and interest in

Table 2. Descriptive analysis of the items and dimensions.

	Total sample (N = 301)
Professional impact of AI	3.62 (0.72)
(1) I believe that the use of AI in my specialty could improve patient care.	3.78 (0.95)
(2) I believe that the use of AI in my specialty could improve clinical decision making.	3.87 (0.87)
(3) I believe that AI can improve population health indicators.	3.81 (0.86)
(4) I believe that AI may change my role as a healthcare professional in the future.	3.51 (0.98)
(5) I believe that the introduction of AI can reduce the economic cost associated with my work.	3.37 (0.97)
(6) I believe that one day AI will be able to be part of my job as a healthcare professional.	3.40 (1.10)
Preparedness for AI	2.76 (0.70)
(7) I believe that most healthcare professionals are prepared for the introduction of AI technology.	2.65 (1.07)
(8) I believe that I have been adequately trained to use AI in my job.	2.02 (1.01)
(9) I believe there is an ethical framework for the use of AI technology in my workplace.	3.23 (1.05)
(10) I believe that if the AI technology makes a mistake, all the responsibility lies with the healthcare professional.	3.15 (1.14)

Mean of scores and standard deviation (SD); m (SD).

AI: artificial intelligence.

AI, answering affirmatively that they understood the concept of AI, would like to be trained in AI, or had received some type of training in AI. Therefore, these results may suggest that receiving training may increase sensitivity to how AI can help in their clinical practice. Otherwise, with regard also to the professional's perceptions, over half of professionals considered that they were prepared for AI (scored 2.76 out of 5 in the Likert scale). This result is similar to the one reported in the Shinnars et al.'s²² study that reported a score on the Likert scale of 2.55 out of 5. It has been seen that the nursing category believes it is better prepared for the application of AI than the medical category. It is true that this category is one of the groups that can generate the most data in healthcare systems, as they play a fundamental role in data collection.⁴⁸ In this context, data collection can become key in the creation of AI algorithms, and it is therefore interesting to take it into account when establishing strategies for adopting AI in the Catalan healthcare system.

Furthermore, it has also been observed that professionals have felt that using AI makes them feel more prepared. Although it has been mentioned that there is currently no integrated AI model in the management programme of the Catalan healthcare system, work is being done on the development of different applications, and there is therefore some awareness of the concept of AI among professionals.

Sheng et al. also show that previous experiences with applied health technology and adequate support for professionals make them more interested in using it and training.⁴⁹

Interestingly, professionals in the Central Catalonia region have considered AI as an ally for improving the quality of care they can provide and have not been so concerned about the possibility of being replaced in the future by AI. This could be due to the long tradition of e-health innovation in this region, being one of the pioneers in the use and application of telemedicine.⁵⁰ In contrast are the responses from a survey of 3018 medical students from 103 medical schools in Turkey, where the results showed that 58.6% agreed that AI would devalue the medical profession and 45.5% were concerned that the application of AI would damage the fundamental value of trust in medicine.⁴³

It should be mentioned that the region of Central Catalonia is currently working to validate an AI algorithm for chest X-rays.³⁸ This region is unique in that it has a shortage of radiologists, which means that PC physicians are increasingly responsible for interpreting X-rays. Given this context, the development of AI algorithms to support X-ray diagnosis, as well as the knowledge and acceptance of healthcare professionals, is essential to integrate and improve workflow for professionals and enhance patient

Table 3. Bivariate analysis between dimensions and sociodemographic variables.

	Professional impact of AI	p-value	Preparedness for AI	p-value
Categorised age (<i>N</i> = 292)		0.628 ^α		0.053 ^α
[18-30]	3.74 (0.59)		3.16 (0.66)	
[30-40]	3.63 (0.82)		2.67 (0.86)	
[40-50]	3.54 (0.69)		2.72 (0.66)	
[50-60]	3.65 (0.75)		2.74 (0.64)	
>= 60	3.70 (0.57)		2.86 (0.59)	
What is your gender?		0.205 ^β		0.435 ^β
Female	3.59 (0.67)		2.74 (0.66)	
Male	3.75 (0.89)		2.84 (0.84)	
What health discipline do you work in?		0.427 ^α		0.024 ^{α,2}
Others	3.86 (1.24)		2.93 (1.11)	
Nursing	3.64 (0.69)		2.86 (0.69)	
Medicine	3.59 (0.69)		2.65 (0.66)	
How many years of experience do you have in your job?		0.758 ^α		0.969 ^α
Less than 1 year	3.52 (0.96)		3.0 (0.85)	
Between 1 and 2 years	3.68 (0.91)		2.94 (0.69)	
Between 3 and 4 years	3.77 (0.63)		2.81 (0.76)	
Between 5 and 10 years	3.52 (0.84)		2.76 (0.81)	
More than 10 years	3.61 (0.67)		2.71 (0.66)	
Do you understand what the concept of AI means?		0.032 ^β		0.128 ^β
No	3.41 (0.69)		2.91 (0.69)	
Yes	3.66 (0.72)		2.74 (0.70)	
In your work do you use AI to perform or plan care?		0.194 ^α		<0.001 ^{α,3}
No	3.60 (0.82)		2.49 (0.66)	
I do not know	3.51 (0.63)		2.80 (0.64)	
Yes	3.70 (0.68)		2.94 (0.71)	
AI education or training:		0.012 ^{α,1}		0.349 ^α
None	3.54 (0.72)		2.71 (0.67)	

(continued)

Table 3. Continued.

	Professional impact of AI	p-value	Preparedness for AI	p-value
Self-starting online courses, webinars, lectures, conferences	3.86 (0.63)		2.80 (0.72)	
Job training	3.74 (0.56)		2.96 (0.77)	
University education	3.45 (1.27)		2.80 (0.86)	
Would you like to receive education on AI in health?		0.006 ^β		0.544 ^β
No	3.21 (0.75)		2.68 (0.68)	
Yes	3.66 (0.70)		2.77 (0.70)	

^αp-Value of ANOVA contrast.

^βp-Value of the T student contrast.

¹Post hoc comparison: only *none* vs *self-starting online courses, webinars, lectures, conferences* has obtained a significant p-value (p = 0.011).

²Post hoc comparisons: only *nursing* vs *medicine* has obtained a significant p-value (p = 0.028).

³Post hoc comparisons: only I don't know vs no (p = 0.006) and yes vs no (p < 0.001) has obtained a significant p-value.

AI: artificial intelligence; ANOVA: analysis of variance.

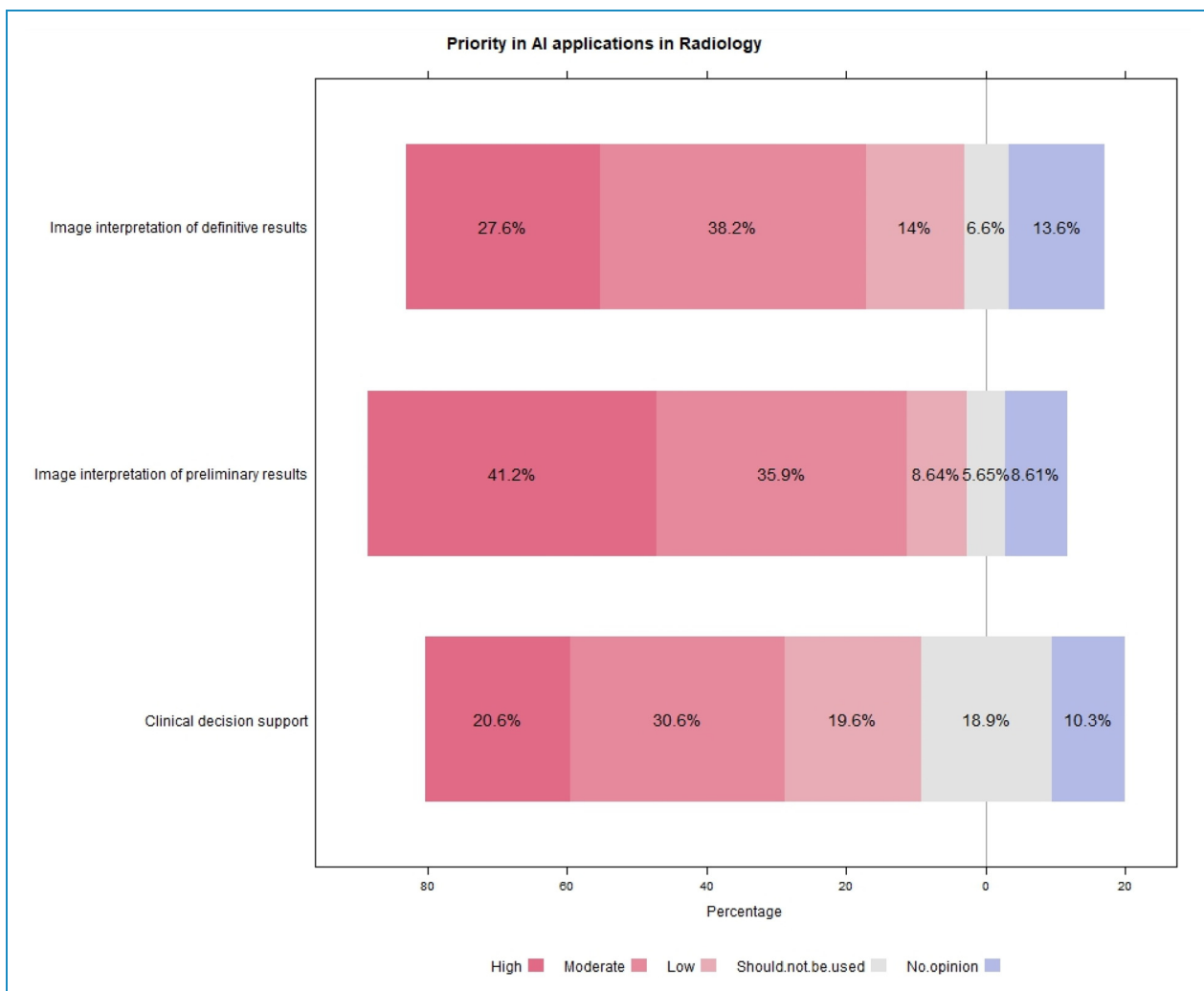


Figure 3. Descriptive analysis of priority in radiology artificial intelligence applications.

safety. In this context, a question was asked about the priority of incorporating different forms of AI in radiology. It should be taken into account that several studies have mentioned a certain reluctance to adopt AI in radiology due to the fact that in the future there will be a decrease in the demand for radiologists to interpret radiographs, as well as the definitive replacement of radiologists in the future.^{29,51} However, both in our study and in other articles in the literature it has been observed that health professionals see clearly that AI should not give a definitive diagnosis to the patient but should only be a tool for quick help and support.^{52,53}

In this context, it seems that professionals in the territory understand AI as an ally to support diagnosis, rather than as a threat, and therefore strategies for the widespread implementation of AI in this territory should be aimed more at training and infrastructures than at resistance to change.

Limitations and strengths

It is possible that there is a bias in the representativeness of the sample, since 81.1% of the professionals who responded to the survey were women.⁵⁴ It should be taken into account, however, that this study was carried out in the PC setting of the Catalan Health Institute of Central Catalonia, where 89.55% of the nursing category and 65.72% of the family medicine category are women.⁵⁵

Furthermore, this is a study with little external validity of the results given the particularities of the territory, which is mostly made up of rural areas. However, this fact brings strength and innovation, as there are few studies on the perception and knowledge of AI in rural areas.

Finally, the SHAIIP tool has only been validated in Australian healthcare settings. To control this limitation, the questionnaire was validated with the study sample and language using Cronbach's alpha and factor analysis.

Conclusion

The results of this study show that most professionals indicated that they had a good understanding of the concept, considered AI to have a positive impact, and felt prepared for its implementation. Additionally, although only as a diagnostic aid, the implementation of AI in radiology was considered a high priority by professionals. Observational studies in selected regions are relevant in order to create strategies that can offer training to professionals to increase their acceptance, reduce resistance to change and prepare them for the near future.

Acknowledgements: The authors would like to thank all the medical and nursing professionals of the Territorial Management of Central Catalonia of the Catalan Institute of Health for their participation in the survey. In addition, this study has been carried out as part of the Industrials Doctorats programme of Catalonia.

Declaration of conflicting interests: The author(s)

declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Ethical approval: The study protocol was approved by the University Institute for Primary Care Research (IDIAP) Jordi Gol Health Care Ethics Committee (Code 20/177-PCV). The survey was anonymous, so the patient consent was not necessary.

Funding: The author(s) received no financial support for the research, authorship, and/or publication of this article.

ORCID iD: Jordi Solé-Casals  <https://orcid.org/0000-0002-6534-1979>

Queralt Miró Catalina  <https://orcid.org/0000-0001-7379-4671>

References

- Shelmerdine SC, Rosendahl K and Arthurs OJ. Artificial intelligence in paediatric radiology: international survey of health care professionals' opinions. *Pediatr Radiol* 2022; 52: 30–41.
- Ramesh AN, Kambhampati C, Monson JRT, et al. Artificial intelligence in medicine. *Ann R Coll Surg Engl* 2004; 86: 334–338.
- Ávila-Tomás JF, Mayer-Pujadas MA and Quesada-Varela VJ. Artificial intelligence and its applications in medicine II: current importance and practical applications. *Aten primaria* 2021; 53: 81–88.
- Karimi A and HaddadPajouh H. Artificial intelligence, important assistant of scientists and physicians. *Galen Med J*. 2020; 9: e2048.
- Tomašev N, Glorot X, Rae JW, et al. A clinically applicable approach to continuous prediction of future acute kidney injury. *Nature* 2019; 572: 116–119.
- Fernandes M, Vieira SM, Leite F, et al. Clinical decision support systems for triage in the emergency department using intelligent systems: a review. *Artif Intell Med* 2020; 102.
- Goldstein A and Shahar Y. An automated knowledge-based textual summarization system for longitudinal, multivariate clinical data. *J Biomed Inform* 2016; 61: 159–175.
- Solé-Casals J, Anchustegui-Echearte I, Martí-Puig P, et al. Discrete cosine transform for the analysis of essential tremor. *Front Physiol* 2019; 9: 1947.
- Ting DSW, Pasquale LR, Peng L, et al. Artificial intelligence and deep learning in ophthalmology. *Br J Ophthalmol* 2019; 103: 167–175.
- Hogarty DT, Su JC, Phan K, et al. Artificial intelligence in dermatology—where we are and the way to the future: a review. *Am J Clin Dermatol* 2020; 21: 41–47.
- Duan F, Huang Z, Sun Z, et al. Topological network analysis of early Alzheimer's disease based on resting-state EEG. *IEEE Trans Neural Syst Rehabil Eng* 2020; 28: 2164–2172.
- Dutta S, Long WJ, Brown DFM, et al. Automated detection using natural language processing of radiologists recommendations for additional imaging of incidental findings. *Ann Emerg Med* 2013; 62: 162–169.

13. Johnson AEW, Pollard TJ, Berkowitz SJ, et al. MIMIC-CXR, a de-identified publicly available database of chest radiographs with free-text reports. *Sci Data* 2019; 6: 1–8.
14. Do HM, Spear LG, Nikpanah M, et al. Augmented radiologist workflow improves report value and saves time: a potential model for implementation of artificial intelligence. *Acad Radiol* 2020; 27: 96–105.
15. Hwang EJ, Park S, Jin K-N, et al. Development and validation of a deep learning-based automated detection algorithm for major thoracic diseases on chest radiographs. *JAMA Netw Open* 2019; 2: e191095.
16. Wu JT, Wong KCL, Gur Y, et al. Comparison of chest radiograph interpretations by artificial intelligence algorithm vs radiology residents. *JAMA Netw Open* 2020; 3: e2022779.
17. Gatt ME, Spectre G, Paltiel O, et al. Chest radiographs in the emergency department: is the radiologist really necessary? *Postgrad Med J* 2003; 79: 214–217.
18. Preston CA, Marr 3rd JJ, Amaraneni KK, et al. Reduction of “callbacks” to the ED due to discrepancies in plain radiograph interpretation. *Am J Emerg Med* 1998; 16: 160–162.
19. Espinosa JA and Nolan TW. Reducing errors made by emergency physicians in interpreting radiographs: longitudinal study. *Br Med J* 2000; 320: 737–740.
20. Martínez-Torres MD, Toral Marín SL, García FB, et al. A technological acceptance of e-learning tools used in practical and laboratory teaching, according to the European higher education area. *Behav Inf Technol* 2008; 27: 495–505.
21. Stevenson JE, Nilsson GC, Petersson GI, et al. Nurses’ experience of using electronic patient records in everyday practice in acute/inpatient ward settings: a literature review. *Health Informatics J* 2010; 16: 63–72.
22. Shinnars L, Aggar C, Grace S, et al. Exploring healthcare professionals’ perceptions of artificial intelligence: validating a questionnaire using the e-Delphi method. *Digit Heal* 2021; 7: 20552076211003430.
23. Nadarzynski T, Miles O, Cowie A, et al. Acceptability of artificial intelligence (AI)-led chatbot services in healthcare: a mixed-methods study. *Digit Heal* 2019; 5: 1–12.
24. Fan W, Liu J, Zhu S, et al. Investigating the impacting factors for the healthcare professionals to adopt artificial intelligence-based medical diagnosis support system (AIMDSS). *Ann Oper Res* 2020; 294: 567–592.
25. Oh S, Kim JH, Choi SW, et al. Physician confidence in artificial intelligence: an online mobile survey. *J Med Internet Res* 2019; 21: e12422.
26. Sarwar S, Dent A, Faust K, et al. Physician perspectives on integration of artificial intelligence into diagnostic pathology. *NPJ Digit Med* 2019; 2: 1–7.
27. Diaz O, Guidi G, Ivashchenko O, et al. Artificial intelligence in the medical physics community: an international survey. *Phys Med* 2021; 81: 141–146.
28. Shinnars L, Aggar C, Grace S, et al. Exploring healthcare professionals’ understanding and experiences of artificial intelligence technology use in the delivery of healthcare: an integrative review. *Health Informatics J* 2020; 26: 1225–1236.
29. Sit C, Srinivasan R, Amlani A, et al. Attitudes and perceptions of UK medical students towards artificial intelligence and radiology: a multicentre survey. *Insights Imaging* 2020; 11: 14.
30. Health Education England. The Topol review: preparing the healthcare workforce to deliver the digital future. An independent report on behalf of the Secretary of State for Health and Social Care. Nhs. 2019:102.
31. Fast E and Horvitz E. Long-term trends in the public perception of artificial intelligence. 31st AAAI Conf Artif Intell AAAI 2017. 2016;963–9.
32. Jha S and Topol EJ. Adapting to artificial intelligence: radiologists and pathologists as information specialists. *JAMA* 2016; 316: 2353–2354.
33. Lai MC, Brian M and Mamzer MF. Perceptions of artificial intelligence in healthcare: findings from a qualitative survey study among actors in France. *J Transl Med* 2020; 18: 14.
34. Shinnars L, Grace S, Smith S, et al. Exploring healthcare professionals’ perceptions of artificial intelligence: piloting the Shinnars Artificial Intelligence Perception tool. *Digit Heal* 2022; 8: 1–8.
35. Peterson RA and Kim Y. On the relationship between coefficient alpha and composite reliability. *J Appl Psychol* 2013; 98: 194–198.
36. Ware JE and Gandek B. Methods for testing data quality, scaling assumptions, and reliability: the IQOLA project approach. *J Clin Epidemiol* 1998; 51: 945–952.
37. Coma E, Miró Q, Medina M, et al. Association between the reduction of face-to-face appointments and the control of patients with type 2 diabetes mellitus during the Covid-19 pandemic in Catalonia. *Diabetes Res Clin Pract* 2021; 182: 109127.
38. Catalina QM, Fuster-Casanovas A, Solé-Casals J, et al. Developing an artificial intelligence model for reading chest X-rays: protocol for a prospective validation study. *JMIR Res Protoc* 2022; 11: e39536.
39. Escalé-Besa A, Fuster-Casanovas A, Börve A, et al. Using artificial intelligence as a diagnostic decision support tool in skin disease: protocol for an observational prospective cohort study. *JMIR Res Protoc* 2022; 11: e37531.
40. Vidal-Alaball J, Fibla DR, Zapata MA, et al. Artificial intelligence for the detection of diabetic retinopathy in primary care: protocol for algorithm development. *JMIR Res Protoc* 2019; 8(2): e12539.
41. Programa per la promoció i desenvolupament de la Intel·ligència Artificial al Sistema de Salut de Catalunya IA.
42. Chen M, Zhang B, Cai Z, et al. Acceptance of clinical artificial intelligence among physicians and medical students: a systematic review with cross-sectional survey. *Front Med* 2022; 9: 990604.
43. Civaner MM, Uncu Y, Bulut F, et al. Artificial intelligence in medical education: a cross-sectional needs assessment. *BMC Med Educ* 2022; 22: 1–9.
44. Wartman SA and Donald Combs C. Medical education must move from the information age to the age of artificial intelligence. *Acad Med* 2018; 93: 1107–1109.
45. Liaw W and Kakadiaris IA. Artificial intelligence and family medicine: better together. *Fam Med* 2020; 52: 8–10.
46. Katznelson G and Gerke S. The need for health AI ethics in medical school education. *Adv Health Sci Educ Theory Pract* 2021; 26: 1447–1458.
47. Lee J, Wu AS, Li D, et al. Artificial intelligence in undergraduate medical education: a scoping review. *Acad Med* 2021; 96: S62–S70.
48. Collins S, Couture B, Kang MJ, et al. Quantifying and visualizing nursing flowsheet documentation burden in acute and critical care. *AMIA Annu Symp Proc* 2018; 2018: 348–357.
49. Liu Sheng OR, Hu PJH and Wei CP, et al. Adoption and diffusion of telemedicine technology in health care organizations: a comparative case study in Hong Kong. *Journal of*

- Organizational Computing and Electronic Commerce*, 2009; 8(4): 247–275.
50. Vidal-Alaball J, Seguí FL, Domingo JLG, et al. Primary care professionals' acceptance of medical record-based, store and forward provider-to-provider telemedicine in Catalonia: results of a web-based survey. *Int J Environ Res Public Health* 2020; 17: 1–13.
 51. Kurowecki D, Lee SY, Monteiro S, et al. Resident physicians' perceptions of diagnostic radiology and the declining interest in the specialty. *Acad Radiol* 2021; 28: 261–270.
 52. Pinto Dos Santos D, Giese D, Brodehl S, et al. Medical students' attitude towards artificial intelligence: a multicentre survey. *Eur Radiol* 2019; 29: 1640–1646.
 53. Akudjedu TN, Torre S, Khine R, et al. Knowledge, perceptions, and expectations of artificial intelligence in radiography practice: a global radiography workforce survey. *J Med Imaging Radiat Sci* 2022; 54: 104–116.
 54. Lundeberg MA, Fox PW and Punçochaí J. Highly confident but wrong: gender differences and similarities in confidence judgments. *J Educ Psychol* 1994; 86: 114–121.
 55. Institut Català de la Salut. L'ICS, una empresa femenina. Institut Català de la Salut [Internet]. 2017 [cited 2023 Apr 27]. Available from: https://ics.gencat.cat/es/detall/noticia/infografia_feminitzacio_ics.html.
-

5.3 Article 3: Knowledge and Perception of the Use of AI and its Implementation in the Field of Radiology: Cross-Sectional Study

Original Paper

Knowledge and Perception of the Use of AI and its Implementation in the Field of Radiology: Cross-Sectional Study

Queralt Miró Catalina^{1,2}, MSc; Joaquim Femenia³, BSc; Aina Fuster-Casanovas¹, MSc; Francesc X Marin-Gomez^{1,2}, PhD; Anna Escalé-Besa^{1,2,3}, MSc; Jordi Solé-Casals^{4,5}, PhD; Josep Vidal-Alaball^{1,2,3}, PhD

¹Unitat de Suport a la Recerca de la Catalunya Central, Fundació Institut Universitari per a la Recerca a l'Atenció Primària de Salut Jordi Gol i Gurina, Sant Fruitós de Bages, Spain

²Health Promotion in Rural Areas Research Group, Gerència Territorial de la Catalunya Central, Institut Català de la Salut, Sant Fruitós de Bages, Spain

³Faculty of Medicine, University of Vic-Central University of Catalonia, Vic, Spain

⁴Data and Signal Processing group, Faculty of Science, Technology and Engineering, University of Vic-Central University of Catalonia, Vic, Spain

⁵Department of Psychiatry, University of Cambridge, Cambridge, United Kingdom

Corresponding Author:

Jordi Solé-Casals, PhD

Data and Signal Processing group

Faculty of Science, Technology and Engineering

University of Vic-Central University of Catalonia

Carrer de la Laura, 13

Vic, 08500

Spain

Phone: 34 938 86 12 22

Email: jordi.sole@uvic.cat

Abstract

Background: Artificial Intelligence (AI) has been developing for decades, but in recent years its use in the field of health care has experienced an exponential increase. Currently, there is little doubt that these tools have transformed clinical practice. Therefore, it is important to know how the population perceives its implementation to be able to propose strategies for acceptance and implementation and to improve or prevent problems arising from future applications.

Objective: This study aims to describe the population's perception and knowledge of the use of AI as a health support tool and its application to radiology through a validated questionnaire, in order to develop strategies aimed at increasing acceptance of AI use, reducing possible resistance to change and identifying possible sociodemographic factors related to perception and knowledge.

Methods: A cross-sectional observational study was conducted using an anonymous and voluntarily validated questionnaire aimed at the entire population of Catalonia aged 18 years or older. The survey addresses 4 dimensions defined to describe users' perception of the use of AI in radiology, (1) "distrust and accountability," (2) "personal interaction," (3) "efficiency," and (4) "being informed," all with questions in a Likert scale format. Results closer to 5 refer to a negative perception of the use of AI, while results closer to 1 express a positive perception. Univariate and bivariate analyses were performed to assess possible associations between the 4 dimensions and sociodemographic characteristics.

Results: A total of 379 users responded to the survey, with an average age of 43.9 (SD 17.52) years and 59.8% (n=226) of them identified as female. In addition, 89.8% (n=335) of respondents indicated that they understood the concept of AI. Of the 4 dimensions analyzed, "distrust and accountability" obtained a mean score of 3.37 (SD 0.53), "personal interaction" obtained a mean score of 4.37 (SD 0.60), "efficiency" obtained a mean score of 3.06 (SD 0.73) and "being informed" obtained a mean score of 3.67 (SD 0.57). In relation to the "distrust and accountability" dimension, women, people older than 65 years, the group with university studies, and the population that indicated not understanding the AI concept had significantly more distrust in the use of AI. On the dimension of "being informed," it was observed that the group with university studies rated access to information more positively and those who indicated not understanding the concept of AI rated it more negatively.

Conclusions: The majority of the sample investigated reported being familiar with the concept of AI, with varying degrees of acceptance of its implementation in radiology. It is clear that the most conflictive dimension is "personal interaction," whereas "efficiency" is where there is the greatest acceptance, being the dimension in which there are the best expectations for the implementation of AI in radiology.

KEYWORDS

artificial intelligence; perception; knowledge; survey; digital health; radiology; public health

Introduction

Artificial Intelligence (AI) has been developing for decades, but its use in the field of health care has experienced an exponential increase in recent years. Currently, there is little doubt that these tools have come to transform clinical practice [1,2]. AI is capable of managing large amounts of information with effectiveness and efficiency beyond the reach of human capability. It is changing clinical care by improving the speed and reliability of diagnostic processes and other health-related procedures [3,4].

Although AI has been used for some time now in some areas of medical processes such as triage support [5], suggesting diagnoses from radiological scans [6] or in specialties such as ophthalmology, dermatology, pathological anatomy, or radiology [7-10], everything suggests that, in a short period of time, these tools will multiply in number and gain weight within the health care field, provided that the ethical and legislative dilemmas raised by their implementation are resolved [11-14].

As for radiology, some tools such as computer-aided diagnosis have been used in the practice of the specialty for decades now. However, with the introduction of new technologies such as deep learning, these tools may become much more powerful and revolutionize this field [15]. This revolution will necessarily have to be accompanied by changes in the training that radiologists receive and in their competencies, but, at the same time, it opens up a new range of opportunities for the specialty [16,17].

In recent years, most studies have focused on the perception of health care professionals regarding the implementation of AI in their practice [18], but it is also necessary to conduct studies focused on the perceptions of users and to consider user preferences to determine their limits and seek the acceptance of society [19].

A study by Ongena et al [20], focused on the field of radiology, showed that patients had little confidence in AI for diagnosis, both in terms of accuracy and confidentiality and especially in terms of personal interaction and communication. In addition, opinions on workflow improvements were ambiguous. However, they preferred AI, as it was able to look at the whole body rather than just specific parts and could report on future diseases.

Furthermore, a qualitative study focused on capturing society's perception of the implementation of AI in health care, in general, showed that most participants agreed that the use of AI could trigger highly beneficial changes and improvements, as well as aid in making diagnoses and treatments much more effective and personalized. That being said, although the overall perception was mostly positive, the implementation of AI also raises concerns about aspects such as privacy [21].

The study conducted in Germany by Fritsch et al [22] showed that there was a good predisposition on the part of the population

to introduce the use of AI in general clinical practice, but that the knowledge of this same population about AI was limited. He also highlighted some demographic groups with more reluctance, including women, elderly people, and people with a low educational level and low technological affinity. To conclude, it evidenced a strong consensus that AI should always be ultimately controlled by a health care professional and that the ultimate responsibility would be that of the health care professional [22].

A study conducted in the United States with 926 participants showed a positive expectation toward the implementation of AI in clinical practice but also revealed some areas in which the implementation of AI raised concerns. They highlighted misdiagnosis, gaps in privacy, or reduced time spent by the physician in their care. Racial and ethnic minority groups were also found to have more concerns [23].

In the modern practice of person-centered health care, it is essential to know what the perception of users is since shared decision-making and patient empowerment are 2 pillars of current health care that have replaced, or will replace in the near future, the doctor-patient paternalism existing in past decades [24].

For all these reasons, this study aims to describe the population's perception and knowledge of the use of AI and its implementation in radiology, through a validated questionnaire, to find out which are the most accepted and problematic areas, and to identify possible sociodemographic factors related in order to develop strategies to increase acceptance and confidence in AI.

Methods

Study Design and Sample

A cross-sectional descriptive study was conducted through a validated, anonymous, and voluntary survey on the use of AI in the radiology setting. The survey was open to any individual in Catalonia who had received the QR code of the survey, or who had visited any Primary Care Centre (CAP) or one of the reference hospitals in the region of Central Catalonia (comprising the counties of l'Anoia, Bages, Berguedà, Osona and Moianès and with population of 525,000 inhabitants). The survey was open to individuals older than 18 years, between September 2022 and March 2023. Responses from individuals residing outside Catalonia were excluded through the postal code of the population of residence.

The survey could be answered in paper format or in digital format, through a QR code that led to the Microsoft 365 questionnaire. Paper sheets and posters with access to the QR were left in all the CAPs of Central Catalonia and in the participating hospital. TeleForm (version 16.5, OpenText Teleform Software) was used to create the survey in paper format and subsequently read the responses.

A minimum of 376 surveys, distributed in the study region were required, to estimate with 95% CI and a precision of 0.08 points, the values of the 4 dimensions of the questionnaire, assuming an SD of 0.75 points [20].

Patient and Public Involvement

The patients and public were not directly involved in the design and conduct of the study due to the cross-sectional nature of the study. It was a survey of the population of Catalonia to know their perception of the implementation of AI. In this context, the population has been the main point of the research and the results reported will be important to establish strategies in the implementation. As per the plan, the study findings will be shared with the administrators of the different sites where the study was conducted to share the results with the general population.

Ethical Considerations

The study protocol was approved by the University Institute for Primary Care Research Jordi Gol Health Care Ethics Committee (Code 20/177-PCV). The survey was completely anonymous and no respondents could be reidentified, so no informed consent was required. It was explained at the beginning of the survey that the data generated would be processed and published. No compensation was paid to those who volunteered to participate in the survey.

Source

A validated questionnaire [20,25] was used to ascertain users' perceptions and knowledge of AI and its use in radiology. Although the original survey contains 5 dimensions, the survey published by the authors only addresses four: "distrust and accountability" (15 questions), "personal interaction" (6 questions), "efficiency" (4 questions), and "being informed" (4 questions), all Likert-type questions (1: strongly disagree, 5: strongly agree). In addition, it contains 5 descriptive Likert-type questions (1: strongly disagree, 5: strongly agree) on the use of computers as a tool in health care.

Permission was requested from the author to use it and translate it from English into Catalan. In order to maintain the fidelity of the original survey, 2 researchers translated it independently, pooled it, and a third helped to reach a consensus in cases of

discrepancies in the translation. With the Likert scoring methodology, results were obtained within a range between 1 and 5. Due to the characteristics of the survey, results closer to 5 refer to a negative perception regarding the use of AI, while results closer to 1 express that this perception is positive.

Additionally, a first sociodemographic part (sex, age, marital status, educational level, and postal code of the population of residence) and 2 questions on knowledge of AI were added. The postal code variable was categorized to obtain the rurality variable. Towns with 10,000 inhabitants or more were considered urban and towns with less than 10,000 inhabitants were considered rural [26].

Statistical Analysis

Categorical variables have been described with absolute frequency and percentage, and continuous variables with mean and SD. In order to calculate dimensions 1-4 of the survey, we took the average of each individual's scale scores for the corresponding questions in each dimension. Cronbach α was used to determine the validity and reliability of 4 dimensions. Typically, Cronbach α of .7 is considered indicative of good internal consistency. However, in some cases, an α of .5 or .6 may still be acceptable [27,28]. To assess the normality of the 4 dimensions, this study used skewness and kurtosis [29-31]. Typically, an absolute skewness value greater than 3 and a kurtosis value greater than 10 may indicate a potential issue with normality. West et al [32] suggested that the absolute value of skewness and kurtosis should not be greater than 2 and 7. For the bivariate analysis between the dimensions and the sociodemographic variables, the Student *t* test or ANOVA with multiple comparisons was used. The analyses were performed with R statistical software (version 4.2.1, R Foundation for Statistical Computing), and the significance level was set at 5%.

Results

A total of 379 people responded to the survey, with 59.8% (n=226) of them being women. The mean age was 43.9 (SD 17.52) years. In addition, 56.9% (n=215) of them had a university education, 51.5% (n=177) of them lived in rural areas, and 89.8% (n=335) of them understood the concept of AI (Table 1).

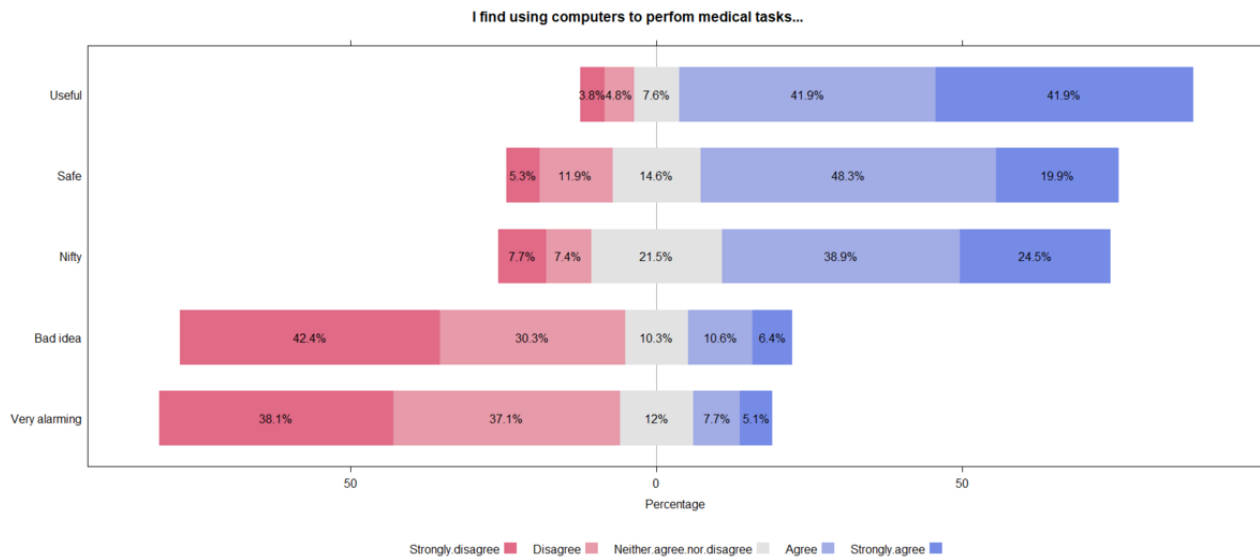
Table 1. Descriptive analysis of the sample (N=379).

Characteristics	Values
Gender (N=378), n (%)	
Female	226 (59.8)
Male	152 (40.2)
Age in years (N=370), mean (SD)	
	43.9 (17.5)
18-34, n (%)	126 (34.1)
35-49, n (%)	102 (27.6)
50-64, n (%)	89 (24.1)
≥65, n (%)	53 (14.2)
Marital status (N=379), n (%)	
Single	147 (38.8)
Married	162 (42.7)
Divorced	24 (6.33)
Widowed	16 (4.22)
Others	30 (7.92)
Educational level (N=378), n (%)	
Does not know or no answer	4 (1.06)
Primary	19 (5.03)
Secondary	34 (8.9)
Baccalaureate, vocational training	106 (28.0)
University students	215 (56.9)
Rurality (N=344), n (%)	
Rural	177 (51.5)
Urban	167 (48.5)
Understand what the concept of artificial intelligence means (N=373), n (%)	
No	38 (10.2)
Yes	335 (89.8)

In relation to the use of computers to perform medical tasks, and considering the negative (strongly disagree and disagree) and positive (strongly agree and agree) options as a single block, [Figure 1](#) shows that 72.7% (n=274) of the population considered that the use of computers to perform medical tasks is not a bad

idea and 68.2% (n=257) thought that it is safe. In addition, 83.8% (n=316) believed that the use of computers is useful for medical tasks and 63.3% (n=239) of the sample thought it would be resourceful, while 75.2% (n=282) did not consider the use of computers in these tasks alarming.

Figure 1. Percentage of opinions on the use of computers to perform medical tasks.



Regarding the 4 dimensions of the survey, and each of their items, the results are expressed on a Likert scale (1: strongly disagree, 5 strongly agree), where 1 reflects positive thinking toward AI and 5 negative thinking. A preliminary analysis was carried out to estimate the internal consistency and reliability of the 4 dimensions through Cronbach α . The dimensions “distrust and accountability” and “personal interaction” obtained an estimate of 0.79 (95% CI 0.75-0.84 and 95% CI 0.75-0.82 respectively), the dimension efficacy an estimate of 0.52 (95% CI 0.42-0.60), and the dimension “being informed” an estimate of 0.42 (95% CI 0.31-0.51).

Table 2 shows the results of the dimensions and items. Following the recommendation, the analysis revealed that the absolute values of skewness and kurtosis for all dimensions were within the acceptable range of <2 and <7 respectively. The mean scores for the different dimensions were 3.37 (SD 0.53) points out of 5 for “distrust and accountability,” 4.37 (SD 0.60) points out of 5 for “personal interaction,” 3.06 (SD 0.73) points out of 5 for “efficiency,” and 3.67 (SD 0.57) points out of 5 for “being informed.”

Table 2. Descriptive analysis of the items and dimensions.

Dimensions	Values, mean (SD)
Dimension 1: Distrust and accountability (N=351). 15 items; Cronbach $\alpha=.79$; (95% CI 0.75-0.84)^a	3.37 (0.53)
AI ^b makes doctors lazy.	2.27 (1.14)
Humans have a better overview than computers on what happens in my body.	3.53 (1.16)
A computer can never compete against the experience of a specialized doctor.	3.44 (1.26)
I think the replacement of doctors by AI will happen in the far future.	2.75 (1.20)
I would never blindly trust a computer.	3.70 (1.23)
AI may prevent errors ^c .	2.13 (0.90)
AI can only be implemented to check human judgment.	3.21 (1.10)
When AI is used, my personal data may fall into the wrong hands.	3.30 (1.27)
I find it worrisome that a computer does not take feelings into account.	3.94 (1.20)
Even if computers are better at evaluating scans, I still prefer a doctor.	3.60 (1.10)
I think radiology is not ready to implement AI in evaluating scans.	2.85 (1.06)
It worries me when computers analyse scans without the interference of humans.	3.87 (1.11)
Through human experience, a radiologist can detect more than a computer.	3.53 (1.08)
It is unclear to me how computers will be used in evaluating scans	3.57 (1.08)
I wonder how it is possible that a computer can give me the results of the scan.	3.10 (1.22)
Dimension 2: Personal interaction (N=364). 6 items; Cronbach $\alpha=.79$; (95% CI 0.75-0.82)^d	4.37 (0.60)
Even when computers are used to evaluate scans, humans always remain responsible.	4.35 (0.89)
As a patient, I want to be treated as a person, not as a number.	4.51 (0.88)
When discussing the results of the scan, humans are indispensable.	4.37 (0.79)
When a computer gives the results, I would miss the explanation.	4.27 (0.94)
Getting the results involves personal contact.	4.10 (1.02)
I find it important to ask questions when getting the results.	4.58 (0.67)
Dimension 3: Efficiency (N=362). 4 items; Cronbach $\alpha=.52$; (95% CI 0.42-0.60)^e	3.06 (0.73)
Evaluating scans with AI will reduce health care waiting times ^c .	2.34 (1.01)
Because of the use of AI, fewer doctors and radiologists are required ^c .	3.51 (1.22)
As far as I am concerned, AI can replace doctors in evaluating scans ^c .	3.06 (1.19)
The sooner I get the results, even when this is from a computer, the more I am at ease.	3.12 (1.15)
Dimension 4: Being informed (N=363). 4 items; Cronbach $\alpha=.42$; (95% CI 0.31-0.51)^f	3.67 (0.57)
When a computer can predict that I will get a disease in the future, I want to know that no matter what.	3.82 (1.16)
If a computer would give the results, I would not feel emotional support.	4.25 (0.96)
A computer should only look at body parts that were selected by my doctor.	3.36 (1.15)
If it does not matter in costs, a computer should always make a full body scan instead of looking at specific body parts.	3.24 (1.35)

^a95% CI 3.32-3.43.

^bAI: artificial intelligence.

^cItems marked are recoded to measure in the same direction.

^d95% CI 4.31-4.43.

^e95% CI 2.98-3.13.

^f95% CI 3.61-3.73.

Finally, the bivariate analysis between the 4 dimensions and the sociodemographic variables is presented in [Table 3](#). In relation to the “distrust and accountability” dimension, women compared to men and people older than 65 years compared to

the other age groups had significantly more distrust in the use of AI ($P=.04$ and $<.01$, respectively). It can also be observed that the group with university studies rated this dimension more positively than the group with baccalaureate studies and

vocational training and that the population that indicated not understanding the AI concept rated it more negatively ($P<.01$ and $.02$ respectively).

Table 3. Bivariate analysis between dimensions and sociodemographic variables.

	Dimension 1		Dimension 2		Dimension 3		Dimension 4	
	Distrust and accountability, mean (SD)	<i>P</i> value	Personal interaction, mean (SD)	<i>P</i> value	Efficiency, mean (SD)	<i>P</i> value	Being informed, mean (SD)	<i>P</i> value
Gender		.04 ^a		.17 ^a		.35 ^a		.95 ^a
Woman	3.29 (0.54)		4.40 (0.62)		3.03 (0.73)		3.67 (0.58)	
Man	3.16 (0.56)		4.31 (0.57)		3.10 (0.74)		3.67 (0.55)	
Age (years)		<.01 ^b		.91 ^b		.05 ^b		.05 ^b
18-34	3.38 (0.50)	a ^c	4.35 (0.58)		3.18 (0.61)		3.62 (0.58)	
35-49	3.34 (0.52)	a	4.41 (0.57)		3.03 (0.79)		3.73 (0.54)	
50-64	3.26 (0.57)	a	4.37 (0.56)		3.02 (0.79)		3.62 (0.62)	
≥65	3.70 (0.45)	b ^c	4.40 (0.76)		2.85 (0.74)		3.85 (0.45)	
Marital status		.31 ^b		.13 ^b		<.01 ^b		.35 ^b
Single	3.38 (0.52)		4.35 (0.61)		3.18 (0.67)	a	3.67 (0.58)	
Married	3.36 (0.55)		4.41 (0.56)		3.04 (0.76)	a	3.71 (0.59)	
Divorced	3.19 (0.51)		4.26 (0.42)		3.13 (0.67)	a	3.47 (0.46)	
Widowed	3.54 (0.47)		4.03 (1.10)		2.18 (0.68)	b	3.55 (0.49)	
Other	3.47 (0.47)		4.51 (0.49)		2.93 (0.63)	a	3.71 (0.50)	
Level of education		<.01 ^b		.59 ^b		.48 ^b		.01 ^b
Does not know or no answer	3.87 (0.42)	ab ^c	4.75 (0.32)		2.62 (0.48)		4.31 (0.37)	a
Primary education	3.55 (0.50)	ab	4.21 (0.49)		2.86 (0.66)		3.71 (0.53)	ab
Secondary education	3.56 (0.58)	ab	4.35 (0.76)		3.04 (0.98)		3.75 (0.55)	ab
Baccalaureate, vocational training	3.46 (0.51)	b	4.38 (0.69)		3.02 (0.74)		3.77 (0.56)	a
University education	3.28 (0.51)	a	4.38 (0.54)		3.10 (0.70)		3.60 (0.57)	b
Residence		.05 ^a		.27 ^a		.34 ^a		.24 ^a
Rural	3.43 (0.49)		4.40 (0.59)		3.01 (0.77)		3.70 (0.55)	
Urban	3.32 (0.56)		4.33 (0.64)		3.09 (0.74)		3.63 (0.59)	
Do you understand the concept of AI?		.02 ^a		.86 ^a		.47 ^a		.04 ^a
No	3.56 (0.48)		4.35 (0.56)		3.14 (0.67)		3.85 (0.53)	
Yes	3.35 (0.53)		4.37 (0.61)		3.05 (0.74)		3.65 (0.57)	

^a*P* value of *t* test.

^b*P* value of ANOVA.

^c“a”, “b”, and “ab”: Different letters indicate significant differences between groups, and groups with the same letter indicate that there are no significant differences between them. For example, for the association between dimension 1 and age, individuals aged 65 years or older had a significantly greater score as compared to individuals aged 18-34, 35-49, and 50-64 years, but there is no difference between the individuals aged 18-34, 35-49, and 50-64 years.

Regarding the dimension of “personal interaction,” there were no significant differences between the demographic characteristics analyzed, and with respect to the dimension of “efficiency” there were only differences according to marital

status, with widowed users showing greater consideration of “efficiency” with respect to the rest, these being the group with the highest mean age ($P<.01$). Finally, on the dimension of “being informed,” it was observed that the group with university

studies rated access to information more positively ($P=.01$) and those who indicated not understanding the concept of AI rated it more negatively ($P=.04$).

Discussion

Principal Findings

This study aimed to describe the level of knowledge and perception, in the population of Catalonia, of the use of AI as a health tool and its implementation in radiology. Of the 4 dimensions analyzed, “distrust and accountability” obtained a mean score of 3.37 (SD 0.53), “personal interaction” obtained a mean score of 4.37 (SD 0.60), “efficiency” obtained a mean score of 3.06 (SD 0.73), and “being informed” obtained a mean score of 3.67 (SD 0.57). In this context, the results obtained provide information on the knowledge and perception of the population and make it possible to find out which are the most problematic areas and which are the most accepted, to develop strategies to increase acceptance of the use of AI.

AI is proving to be a tool that will become fundamental in many aspects of people’s future lives and also in health care practice. In the field of diagnostic imaging, this evolution is particularly rapid and is likely to generate ethical, legal, and social conflicts over its use and acceptance [33-35]. Although patient autonomy must always be respected and any action should be individualized, knowing the population’s overall perception of the matter could help to place the patient at the center of health care. It will also be important to educate and raise awareness among both health care professionals and the general population and, in order to make these training or awareness programs more efficient, it is necessary to know in which areas there is greater distrust.

The results show that there is a high percentage of the population analyzed that has notions about the concept of AI. It must be assumed that this percentage will continue to increase since this technology is being introduced in more and more areas and is opening up to the general public, who can now make use of some of these web-based tools. The results obtained suggest an inherent resistance to the use of AI in the field of radiology, since in the 4 dimensions analyzed, a more negative assessment was obtained. Specifically, “personal interaction” was the most negatively rated dimension, while “efficiency” was the dimension in which the population analyzed was most confident.

These results are similar to those of the study conducted by Ongena et al [20] in the Netherlands and suggest that the population believes that the use of AI can improve and reduce waiting time in their medical care. However, it still raises quite a few doubts about the fact that their health care is not supervised by a human, as well as about the need for human interaction in the medical process derived from radiological studies.

Specifically, for the most negatively rated dimension, which was “personal interaction,” the results may suggest that a large part of the distrust generated by the implementation of AI in diagnostic imaging is due to depersonalization, feeling that you have not received the care you need or that the medical professional has not devoted the necessary time to your case.

It is noteworthy that this fear is much more intense than the doubts that a diagnosis made by AI can cause, as evidenced by the questions “I would never blindly trust a computer,” which scores 3.70 out of 5, or “It worries me when computers analyze scans without interference of humans,” which scores a 3.87 out of 5. Moreover, this value in personal interaction remains constant across all sociodemographic groups and, therefore, reveals itself as a focal point in medical care.

Richardson et al [36] conducted 15 focus groups with adult patients who had recently visited primary care centers in order to analyze the emergence of attitudes and beliefs about health care AI. After analyzing the results, the authors proposed a conceptual framework for understanding patient attitudes and beliefs about health care AI. The attitudes and beliefs about AI used in health care are initially shaped by the patient’s past experiences. Previous illness, the use of technology in health care, the relationship between health care providers, the comfort of the patient using the technology, as well as the wider social context of the person are the main themes highlighted by patients. All of these experiences contribute to shaping the patient’s beliefs about health care and technology, which ultimately influences the development of their particular attitude toward health care AI. In this context, predicting how patients will develop an attitude toward AI in health care becomes crucial for its successful implementation.

With respect to the sociodemographic characteristics of the sample, the results obtained are noteworthy. Women and the population older than 65 years have a more negative view of the “distrust and accountability” dimension, while users with a university education have a less conflictive view of this dimension. These results are also observed in Fritsch et al’s [22] study and can be explained by the fact that the proportion of university students is likely to be lower in the 65-year-old age group. It can also be inferred that a higher level of education correlates with a higher degree of understanding and use of new technologies, which would increase confidence in them.

It has also been observed that university students are the demographic group that most positively values access to information, following the line of the study [37]. This could be linked to the fact that this more educated population group feels more capable of making and evaluating their decisions.

Furthermore, it was observed that users who indicated that they were aware of the concept of AI more positively rated access to information and the impact of AI on “distrust and accountability.” Therefore, these results may suggest that being trained or having received training may increase sensitivity to how AI can be beneficial in the health domain.

For all these reasons, the implementation of AI in the field of radiology appears to be an inexorable reality, but it must necessarily go hand in hand with acceptance by the general population taking into account cultural aspects and prior knowledge and perceptions. Studies such as the one carried out are important to take the pulse of society and design strategies to ensure that this evolution takes place under an umbrella of acceptance. Leaving users out of this process would be a mistake that could have ethical and legal consequences that we can only now begin to anticipate.

Limitations and Strengths

As limitations of the study, we note an age bias, since the concepts used in the survey may be unfamiliar to older people. Due to this selection bias, there is more representation of younger age groups in the sample. Considering that the survey could be easily distributed and accessed via a QR code and that the subject was especially attractive among the younger population, these circumstances might have led to a higher representation of young individuals in the study sample, also resulting in a higher representation of the individuals with university level education. There could also be a selection bias due to the fact that the study population was people who had attended a health center during the study period. In addition, the survey was only translated into Catalan, the predominant language of the region studied, but with 2 official languages, Catalan and Spanish. This might have limited participation from individuals who primarily speak Spanish. Finally, there could be duplicate responses by the same user, given the web-based survey formats. While we made efforts to minimize this by using unique identifiers and tracking IP addresses, we cannot entirely rule out the chance of duplicate or fraudulent responses. This is a limitation inherent to web-based surveys. As strengths,

we found a good rural or urban representation and a high pervasiveness of the AI concept that gives solidity to the results obtained.

Conclusions

The results of the study show that the majority of the population reported being familiar with the concept of AI, with varying degrees of acceptance of its implementation in radiology. It is clear that the dimension where the population has shown the most disagreement has been “personal interaction,” while in the field of “efficiency” is where there is greater acceptance, being the dimension in which there are better expectations regarding the implementation of AI in radiology. These findings underscore the importance of considering cultural aspects, public perceptions, and knowledge when implementing AI in health care, with a focus on addressing concerns related to depersonalization and ensuring a balance between technological advancement and human interaction. This study may be helpful in creating strategies, depending on the profile of the population, to increase acceptance, reduce resistance to change, and prepare the population for a future where AI will be more and more present in health care.

Acknowledgments

The authors would like to thank all those who responded to the survey. The authors would also like to thank the professionals of the Territorial Management of Central Catalonia of the Catalan Institute of Health who have disseminated the survey and helped the population to answer it. This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors. The authors have not received any funding to develop the paper.

Data Availability

The data sets generated and analyzed during this study are not publicly available because the paper was based on confidential data but are available from the corresponding author upon reasonable request.

Authors' Contributions

All authors were responsible for the conceptualization and design of this study and contributed to the final paper. QMC and JF recruited participants and collected data, performed the statistical analysis and interpretation of the data, and drafted the paper. AFC participated in the drafted paper and critical revision of the paper. AEB, FXMG, JSC, and JVA reviewed and revised the paper.

Conflicts of Interest

None declared.

References

1. Bélisle-Pipon JC, Couture V, Roy MC, Ganache I, Goetghebeur M, Cohen IG. What makes artificial intelligence exceptional in health technology assessment? *Front Artif Intell*. 2021;4:736697. [FREE Full text] [doi: [10.3389/frai.2021.736697](https://doi.org/10.3389/frai.2021.736697)] [Medline: [34796318](https://pubmed.ncbi.nlm.nih.gov/34796318/)]
2. Karimi A, HaddadPajouh H. Artificial intelligence, important assistant of scientists and physicians. *Galen Med J*. 2020;9:e2048. [FREE Full text] [doi: [10.31661/gmj.v9i0.2048](https://doi.org/10.31661/gmj.v9i0.2048)] [Medline: [34466625](https://pubmed.ncbi.nlm.nih.gov/34466625/)]
3. Ramesh AN, Kambhampati C, Monson JRT, Drew PJ. Artificial intelligence in medicine. *Ann R Coll Surg Engl*. 2004;86(5):334-338. [FREE Full text] [doi: [10.1308/147870804290](https://doi.org/10.1308/147870804290)] [Medline: [15333167](https://pubmed.ncbi.nlm.nih.gov/15333167/)]
4. Ávila-Tomás JF, Mayer-Pujadas MA, Quesada-Varela VJ. Artificial intelligence and its applications in medicine II: current importance and practical applications. *Aten Primaria*. 2021;53(1):81-88. [FREE Full text] [doi: [10.1016/j.aprim.2020.04.014](https://doi.org/10.1016/j.aprim.2020.04.014)] [Medline: [32571595](https://pubmed.ncbi.nlm.nih.gov/32571595/)]
5. Tomašev N, Glorot X, Rae JW, Zielinski M, Askham H, Saraiva A, et al. A clinically applicable approach to continuous prediction of future acute kidney injury. *Nature*. 2019;572(7767):116-119. [FREE Full text] [doi: [10.1038/s41586-019-1390-1](https://doi.org/10.1038/s41586-019-1390-1)] [Medline: [31367026](https://pubmed.ncbi.nlm.nih.gov/31367026/)]

6. Goldstein A, Shahar Y. An automated knowledge-based textual summarization system for longitudinal, multivariate clinical data. *J Biomed Inform.* 2016;61:159-175. [[FREE Full text](#)] [doi: [10.1016/j.jbi.2016.03.022](https://doi.org/10.1016/j.jbi.2016.03.022)] [Medline: [27039119](#)]
7. Ting DSW, Pasquale LR, Peng L, Campbell JP, Lee AY, Raman R, et al. Artificial intelligence and deep learning in ophthalmology. *Br J Ophthalmol.* 2019;103(2):167-175. [[FREE Full text](#)] [doi: [10.1136/bjophthalmol-2018-313173](https://doi.org/10.1136/bjophthalmol-2018-313173)] [Medline: [30361278](#)]
8. Hogarty DT, Su JC, Phan K, Attia M, Hossny M, Nahavandi S, et al. Artificial intelligence in dermatology-where we are and the way to the future: a review. *Am J Clin Dermatol.* 2020;21(1):41-47. [[FREE Full text](#)] [doi: [10.1007/s40257-019-00462-6](https://doi.org/10.1007/s40257-019-00462-6)] [Medline: [31278649](#)]
9. Dutta S, Long WJ, Brown DFM, Reisner AT. Automated detection using natural language processing of radiologists recommendations for additional imaging of incidental findings. *Ann Emerg Med.* 2013;62(2):162-169. [[FREE Full text](#)] [doi: [10.1016/j.annemergmed.2013.02.001](https://doi.org/10.1016/j.annemergmed.2013.02.001)] [Medline: [23548405](#)]
10. Niazi MKK, Parwani AV, Gurcan MN. Digital pathology and artificial intelligence. *Lancet Oncol.* 2019;20(5):e253-e261. [[FREE Full text](#)] [doi: [10.1016/S1470-2045\(19\)30154-8](https://doi.org/10.1016/S1470-2045(19)30154-8)] [Medline: [31044723](#)]
11. Amann J, Blasimme A, Vayena E, Frey D, Madai VI, Precise4Q consortium. Explainability for artificial intelligence in healthcare: a multidisciplinary perspective. *BMC Med Inform Decis Mak.* 2020;20(1):310. [[FREE Full text](#)] [doi: [10.1186/s12911-020-01332-6](https://doi.org/10.1186/s12911-020-01332-6)] [Medline: [33256715](#)]
12. Badnjević A, Avdihodžić H, Pokvić LG. Artificial intelligence in medical devices: past, present and future. *Psychiatr Danub.* 2021;33(Suppl 3):S336-S341. [[FREE Full text](#)] [Medline: [34010259](#)]
13. Couture V, Roy MC, Dez E, Laperle S, Bélisle-Pipon JC. Ethical implications of artificial intelligence in population health and the public's role in its governance: perspectives from a citizen and expert panel. *J Med Internet Res.* 2023;25:e44357. [[FREE Full text](#)] [doi: [10.2196/44357](https://doi.org/10.2196/44357)] [Medline: [37104026](#)]
14. Tripathi S, Musiolik TH. Fairness and ethics in artificial intelligence-based medical imaging. In: *Research Anthology on Improving Medical Imaging Techniques for Analysis and Intervention*. Hershey, PA: IGI Global; 2022.
15. Chan HP, Samala RK, Hadjiiski LM. CAD and AI for breast cancer-recent development and challenges. *Br J Radiol.* 2020;93(1108):20190580. [[FREE Full text](#)] [doi: [10.1259/bjr.20190580](https://doi.org/10.1259/bjr.20190580)] [Medline: [31742424](#)]
16. Gorospe-Sarasúa L, Muñoz-Olmedo JM, Sendra-Portero F, de Luis-García R. Challenges of radiology education in the era of artificial intelligence. *Radiologia (Engl Ed).* 2022;64(1):54-59. [[FREE Full text](#)] [doi: [10.1016/j.rxeng.2020.10.012](https://doi.org/10.1016/j.rxeng.2020.10.012)] [Medline: [35180987](#)]
17. Catalina QM, Fuster-Casanovas A, Solé-Casals J, Vidal-Alaball J. Developing an artificial intelligence model for reading chest X-rays: protocol for a prospective validation study. *JMIR Res Protoc.* 2022;11(11):e39536. [[FREE Full text](#)] [doi: [10.2196/39536](https://doi.org/10.2196/39536)] [Medline: [36383419](#)]
18. Catalina QM, Fuster-Casanovas A, Vidal-Alaball J, Escalé-Besa A, Marin-Gomez FX, Femenia J, et al. Knowledge and perception of primary care healthcare professionals on the use of artificial intelligence as a healthcare tool. *Digit Health.* 2023;9:20552076231180511. [[FREE Full text](#)] [doi: [10.1177/20552076231180511](https://doi.org/10.1177/20552076231180511)] [Medline: [37361442](#)]
19. Haan M, Ongena YP, Hommes S, Kwee TC, Yakar D. A qualitative study to understand patient perspective on the use of artificial intelligence in radiology. *J Am Coll Radiol.* 2019;16(10):1416-1419. [doi: [10.1016/j.jacr.2018.12.043](https://doi.org/10.1016/j.jacr.2018.12.043)] [Medline: [30878311](#)]
20. Ongena YP, Haan M, Yakar D, Kwee TC. Patients' views on the implementation of artificial intelligence in radiology: development and validation of a standardized questionnaire. *Eur Radiol.* 2020;30(2):1033-1040. [[FREE Full text](#)] [doi: [10.1007/s00330-019-06486-0](https://doi.org/10.1007/s00330-019-06486-0)] [Medline: [31705254](#)]
21. McCradden MD, Sarker T, Paprica PA. Conditionally positive: a qualitative study of public perceptions about using health data for artificial intelligence research. *BMJ Open.* 2020;10(10):e039798. [[FREE Full text](#)] [doi: [10.1136/bmjopen-2020-039798](https://doi.org/10.1136/bmjopen-2020-039798)] [Medline: [33115901](#)]
22. Fritsch SJ, Blankenheim A, Wahl A, Hetfeld P, Maassen O, Deffge S, et al. Attitudes and perception of artificial intelligence in healthcare: a cross-sectional survey among patients. *Digit Health.* 2022;8:20552076221116772. [[FREE Full text](#)] [doi: [10.1177/20552076221116772](https://doi.org/10.1177/20552076221116772)] [Medline: [35983102](#)]
23. Khullar D, Casalino LP, Qian Y, Lu Y, Krumholz HM, Aneja S. Perspectives of patients about artificial intelligence in health care. *JAMA Netw Open.* 2022;5(5):e2210309. [[FREE Full text](#)] [doi: [10.1001/jamanetworkopen.2022.10309](https://doi.org/10.1001/jamanetworkopen.2022.10309)] [Medline: [35507346](#)]
24. Hickmann E, Richter P, Schlieter H. All together now - patient engagement, patient empowerment, and associated terms in personal healthcare. *BMC Health Serv Res.* 2022;22(1):1116. [[FREE Full text](#)] [doi: [10.1186/s12913-022-08501-5](https://doi.org/10.1186/s12913-022-08501-5)] [Medline: [36056354](#)]
25. Centerdata. Artificial intelligence in medicine. Centerdata. 2021. URL: https://www.dataarchive.lissdata.nl/study_units/view/1089 [accessed 2023-09-28]
26. Coma E, Mora N, Méndez L, Benítez M, Hermosilla E, Fàbregas M, et al. Primary care in the time of COVID-19: monitoring the effect of the pandemic and the lockdown measures on 34 quality of care indicators calculated for 288 primary care practices covering about 6 million people in Catalonia. *BMC Fam Pract.* 2020;21(1):208. [[FREE Full text](#)] [doi: [10.1186/s12875-020-01278-8](https://doi.org/10.1186/s12875-020-01278-8)] [Medline: [33038926](#)]

27. Peterson RA, Kim Y. On the relationship between coefficient alpha and composite reliability. *J Appl Psychol*. 2013;98(1):194-198. [doi: [10.1037/a0030767](https://doi.org/10.1037/a0030767)] [Medline: [23127213](https://pubmed.ncbi.nlm.nih.gov/23127213/)]
28. Ware JE, Gandek B. Methods for testing data quality, scaling assumptions, and reliability: the IQOLA project approach. *J Clin Epidemiol*. 1998;51(11):945-952. [FREE Full text] [doi: [10.1016/s0895-4356\(98\)00085-7](https://doi.org/10.1016/s0895-4356(98)00085-7)] [Medline: [9817111](https://pubmed.ncbi.nlm.nih.gov/9817111/)]
29. Hair JF, Hair J, Black WC, Babin BJ, Anderson RE. *Multivariate Data Analysis*. Seventh Edition. Edinburgh. Pearson Education Limited; 2010;1-758.
30. Kline RB. *Principles and Practice of Structural Equation Modeling*. 5th Edition. New York, NY. Guilford; 2011;3-427.
31. Tabachnick BG, Fidell LS, ProQuest. *Using Multivariate Statistics*. 6th Edition. Harlow. Pearson Education Limited; 2014;1-983.
32. West SG, Finch JF, Curran JF. Structural equation models with nonnormal variables: problems and remedies. In: RH H, editor. *Structural Equation Modeling: Concepts, Issues, and Applications*. Thousand Oaks. SAGE Publications; 1995.
33. Ahmad Z, Rahim S, Zubair M, Abdul-Ghafar J. Artificial Intelligence (AI) in medicine, current applications and future role with special emphasis on its potential and promise in pathology: present and future impact, obstacles including costs and acceptance among pathologists, practical and philosophical considerations. a comprehensive review. *Diagn Pathol*. 2021;16(1):24. [FREE Full text] [doi: [10.1186/s13000-021-01085-4](https://doi.org/10.1186/s13000-021-01085-4)] [Medline: [33731170](https://pubmed.ncbi.nlm.nih.gov/33731170/)]
34. Nakamoto T, Takahashi W, Haga A, Takahashi S, Kiryu S, Nawa K, et al. Prediction of malignant glioma grades using contrast-enhanced T1-weighted and T2-weighted magnetic resonance images based on a radiomic analysis. *Sci Rep*. 2019;9(1):19411. [FREE Full text] [doi: [10.1038/s41598-019-55922-0](https://doi.org/10.1038/s41598-019-55922-0)] [Medline: [31857632](https://pubmed.ncbi.nlm.nih.gov/31857632/)]
35. Liu X, Zhou H, Hu Z, Jin Q, Wang J, Ye B. Clinical application of artificial intelligence recognition technology in the diagnosis of stage T1 lung cancer. *Zhongguo Fei Ai Za Zhi*. 2019;22(5):319-323. [FREE Full text] [doi: [10.3779/j.issn.1009-3419.2019.05.09](https://doi.org/10.3779/j.issn.1009-3419.2019.05.09)] [Medline: [31109442](https://pubmed.ncbi.nlm.nih.gov/31109442/)]
36. Richardson JP, Curtis S, Smith C, Pacyna J, Zhu X, Barry B, et al. A framework for examining patient attitudes regarding applications of artificial intelligence in healthcare. *Digit Health*. 2022;8:20552076221089084. [FREE Full text] [doi: [10.1177/20552076221089084](https://doi.org/10.1177/20552076221089084)] [Medline: [35355806](https://pubmed.ncbi.nlm.nih.gov/35355806/)]
37. De Santis KK, Jahnel T, Sina E, Wienert J, Zeeb H. Digitization and health in Germany: cross-sectional nationwide survey. *JMIR Public Health Surveill*. 2021;7(11):e32951. [FREE Full text] [doi: [10.2196/32951](https://doi.org/10.2196/32951)] [Medline: [34813493](https://pubmed.ncbi.nlm.nih.gov/34813493/)]

Abbreviations

AI: artificial intelligence

CAP: Primary Care Centre

Edited by T de Azevedo Cardoso; submitted 11.07.23; peer-reviewed by J Thrall, V Shalin; comments to author 16.08.23; revised version received 31.08.23; accepted 25.09.23; published 13.10.23

Please cite as:

Miró Catalina Q, Femenia J, Fuster-Casanovas A, Marin-Gomez FX, Escalé-Besa A, Solé-Casals J, Vidal-Alaball J. Knowledge and Perception of the Use of AI and its Implementation in the Field of Radiology: Cross-Sectional Study *J Med Internet Res* 2023;25:e50728

URL: <https://www.jmir.org/2023/1/e50728>

doi: [10.2196/50728](https://doi.org/10.2196/50728)

PMID: [37831495](https://pubmed.ncbi.nlm.nih.gov/37831495/)

©Queralt Miró Catalina, Joaquim Femenia, Aina Fuster-Casanovas, Francesc X Marin-Gomez, Anna Escalé-Besa, Jordi Solé-Casals, Josep Vidal-Alaball. Originally published in the Journal of Medical Internet Research (<https://www.jmir.org>), 13.10.2023. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in the Journal of Medical Internet Research, is properly cited. The complete bibliographic information, a link to the original publication on <https://www.jmir.org/>, as well as this copyright and license information must be included.

5.4 Article 4: Real-world testing of an artificial intelligence algorithm for the analysis of chest X-rays in primary care settings



OPEN

Real-world testing of an artificial intelligence algorithm for the analysis of chest X-rays in primary care settings

Queralt Miró Catalina^{1,2,3}, Josep Vidal-Alaball^{1,2,4}✉, Aïna Fuster-Casanovas^{1,2}, Anna Escalé-Besa^{1,2,4}, Anna Ruiz Comellas^{1,2,4} & Jordi Solé-Casals^{5,6}✉

Interpreting chest X-rays is a complex task, and artificial intelligence algorithms for this purpose are currently being developed. It is important to perform external validations of these algorithms in order to implement them. This study therefore aims to externally validate an AI algorithm's diagnoses in real clinical practice, comparing them to a radiologist's diagnoses. The aim is also to identify diagnoses the algorithm may not have been trained for. A prospective observational study for the external validation of the AI algorithm in a region of Catalonia, comparing the AI algorithm's diagnosis with that of the reference radiologist, considered the gold standard. The external validation was performed with a sample of 278 images and reports, 51.8% of which showed no radiological abnormalities according to the radiologist's report. Analysing the validity of the AI algorithm, the average accuracy was 0.95 (95% CI 0.92; 0.98), the sensitivity was 0.48 (95% CI 0.30; 0.66) and the specificity was 0.98 (95% CI 0.97; 0.99). The conditions where the algorithm was most sensitive were external, upper abdominal and cardiac and/or valvular implants. On the other hand, the conditions where the algorithm was less sensitive were in the mediastinum, vessels and bone. The algorithm has been validated in the primary care setting and has proven to be useful when identifying images with or without conditions. However, in order to be a valuable tool to help and support experts, it requires additional real-world training to enhance its diagnostic capabilities for some of the conditions analysed. Our study emphasizes the need for continuous improvement to ensure the algorithm's effectiveness in primary care.

Radiology is the speciality of medicine that uses different imaging techniques to detect and treat diseases. One of the most widely used imaging techniques in this field is radiography (X-ray), used to generate images of the inside of the body, which has the advantage of being quick to perform and requiring low levels of radiation^{1,2}. Within the scope of X-rays, one of the most common tests is the chest X-ray³⁻⁶, used to detect pulmonary and cardiovascular diseases, among others.

Despite the importance of this speciality and the volume of radiological tests that exist, the lack of radiologists able to interpret them⁷⁻¹¹, as well as the increase in errors in interpretation as a consequence of a heavy workload is becoming increasingly evident¹²⁻¹⁵. Furthermore, radiological diagnosis is a component of clinical competencies in various specialities, including family and community medicine. It is a common practice for practitioners in this field to interpret chest X-rays, despite their lower degree of expertise¹⁶. This fact highlights the need and, therefore, the opportunity to introduce tools such as artificial intelligence (AI) into this field to support radiologists and other healthcare practitioners who need to interpret an X-ray.

¹Unitat de Suport a la Recerca de la Catalunya Central, Fundació Institut Universitari per a la Recerca a l'Atenció Primària de Salut Jordi Gol i Gurina, Sant Fruitós de Bages, Spain. ²Health Promotion in Rural Areas Research Group, Gerència d'Atenció Primària i a la Comunitat de la Catalunya Central, Institut Català de la Salut, Carrer Pica d'Estats, 13-15, 08272 Sant Fruitós de Bages, Barcelona, Spain. ³Faculty of Science Technology and Engineering, University of Vic-Central University of Catalonia, Vic, Spain. ⁴Faculty of Medicine, University of Vic-Central University of Catalonia, Vic, Spain. ⁵Data and Signal Processing Group, Faculty of Science, Technology and Engineering, University of Vic-Central University of Catalonia, Vic, Spain. ⁶Department of Psychiatry, University of Cambridge, Cambridge, UK. ✉email: jvidal.cc.ics@gencat.cat; jordi.sole@uvic.cat

Technological support tools, such as computer-aided diagnostics (CAD), have long been used in this area. However, the advent of deep learning and machine learning models in recent years has offered the potential to develop new support tools aiming to overcome the primary limitations of CAD and enhance accuracy^{17,18}. Deep learning models, compared to CAD, are built to train and work with large databases, to have constant improvement over time by learning from errors, and to have the ability to detect more than one condition at a time, all of which makes them much more powerful than CAD.

In this context, AI algorithms, including deep and machine learning models can assist in diagnosis, potentially enhancing diagnostic accuracy. However, while AI will not replace professionals, it is essential to acknowledge that the implementation of AI in routine clinical practice must be both safe and effective^{20,21}. One of the current concerns lies in the fact that many of the studies on applications of new AI models only present in silico validation, a phenomenon called “digital exceptionalism”, without performing external validation in the actual implementation environment. External validation is important, as it allows estimating the accuracy of the model in a population different from the training population, selected in real clinical practice, thus allowing the subsequent generalisation of the results^{22–26}.

A recent study conducting external validation of an AI algorithm designed to classify chest X-rays as normal or abnormal using a cohort of real images from two primary care centres, indicated that additional training with data from these environments was required to enhance the algorithm’s performance, particularly in clinical setting different from its initial training environment²⁷.

In this context, external validation, crucial for ensuring non-discrimination and equity in healthcare, should be a key requirement for the widespread implementation of AI algorithms. However, it is not yet specifically mandated by European legislation (Regulation 2017/745) and this not a prerequisite for marketing an AI algorithm²⁸. For this reason, different groups of experts around the world have developed guidelines to stipulate the essential requirements for using AI algorithms as a complementary diagnostic tool. Yet, while expert groups emphasise the necessity of external validation to confirm its practical potential in clinical settings, this requirement is not yet mandated in the Regulation^{26,29,30}.

The implementation of AI in healthcare appears to be an imminent reality that can offer significant benefits to both professionals and the general population. However, it is essential to implement safe and validated tools in real clinical settings to maintain fairness. In this sense, this study aims to externally validate an AI algorithm’s diagnoses in real primary care settings, comparing them to a radiologist’s diagnoses. The aim is also to identify diagnoses the algorithm may not have been trained for.

Materials and methods

The study protocol has been previously described and published³². Nevertheless, the most relevant points of the present study are described below.

The ChestEye AI algorithm

Oxipit³³, one of the leading companies in AI medical image reading, has developed a fully automatic computer-aided diagnosis (CAD) AI algorithm for reading chest X-rays trained with more than 300,000 images, available through a web platform called ChestEye. The ChestEye imaging service has been certified as a Class II medical device on the Australian Register of Therapeutic Goods and has also been CE marked³³. The web platform reads the inserted chest X-ray and returns the automatic report with the ability to detect 75 conditions, which cover 90% of diagnoses, as well as a heat map to show the locations of the findings. Thus, ChestEye allows radiologists to analyse only the most relevant X-rays^{33,34}.

Study design

A prospective observational study for the external validation of the AI algorithm in a region of Catalonia of users who were scheduled for chest radiography at the Osona Primary Care Center. For each user, the report of the reference radiologist (considered the gold standard) was obtained. Subsequently, the research team input the image into the AI algorithm to obtain the diagnosis. This allowed for the comparison of the AI’s performance with the reference standard in terms of accuracy, sensitivity, specificity, positive predictive value and negative predictive value c.

Description of the study population, time frame, and data collection

The study was carried out at the Catalan Institute of Health’s Primary Care Centre Vic Nord (Osona, Catalonia, Spain), a reference centre where all chest X-rays in the region are performed (with a coverage of 125,000 users). At this same centre, convenience recruitment was carried out from 7 February 2022 to 31 May 2022. The study was explained, and the information sheet and informed consent were given to all patients who came for a chest X-ray and met the inclusion criteria³².

The reference population of the study was the entire population of the Osona region who underwent chest X-rays at the study centre and agreed to participate in the study. The study included only anteroposterior chest X-rays on those over 18 years of age and excluded pregnant women and poor-quality chest X-rays (poor exposure, non-centred or rotated images).

Sample size

Due to a problem with the image collection centre, the sample size calculated in the protocol³² could not be reached. For this reason, the sample was recalculated by increasing the precision by one percentage point. Thus, in order to validate the algorithm, a sample of 450 images was needed to estimate an overall accuracy expected

to be around 80%, with a confidence interval of 95%, a precision level of 5% and a percentage of replacements needed of 15%.

Procedure

Once recruitment was completed, the Technical Service of the Catalan Institute of Health of Central Catalonia extracted the patients' anonymous and automated images and their corresponding non-anonymised reports. The images and reports were then coded together so that they could be related.

The research team then entered all the images into the AI model to extract their interpretation (the diagnosis or the possibility of no abnormalities). At the same time, three general practitioners interpreted the reference radiologists' reports, without seeing the images in order to avoid assumptions, with the aim of extracting the diagnoses described.

Finally, the group of general practitioners grouped all the conditions detectable by the AI model into 9 categories according to the anatomy of the thorax in order to build an individual and grouped study. The categories were: external implants, mediastinal findings or conditions, cardiac and valvular conditions, vessel conditions, bone conditions, pleural or pleural space conditions, upper abdominal findings or conditions, pulmonary parenchymal findings or conditions, and others.

Statistical analysis

To validate the algorithm, the AI algorithm's diagnoses were compared with those of the gold standard. The accuracy of the algorithm and the confusion matrix were obtained from the images correctly classified positive (PT), correctly classified negative (TN), false positive (FP) and false negative (FN). Sensitivity and specificity were also calculated. These measurements were obtained for the total sample, for each condition and for each of the categories according to physiology. Analyses were performed with R software version 4.2.1 and all confidence intervals were 95%.

Ethics committee

The University Institute for Research in Primary Health Care Jordi Gol i Gurina (Barcelona, Spain) ethics committee approved the trial study protocol (approval code: 21/288). Written informed consent was requested from all patients participating in the study.

Ethical considerations

Radiologists' assessment and decisions were not influenced by this study, as the normal radiology referral workflow was not affected. This project was approved by the Research Ethics Committee (REC) from the Foundation University Institute for Primary Health Care Research Jordi Gol i Gurina (IDIAPJGol) (P21/288-P). The study was performed in accordance with relevant guidelines/regulations, and informed consent was obtained from all participants. All research was performed in accordance with the Declaration of Helsinki.

Results

Of the 471 patients who agreed to participate in the study and provide the images and reports, the final sample for external validation of the model was 278, mainly due to computer-related issues when extracting the data. In some cases, when automatically extracting images and reports, both were not obtained, i.e., either the image was missing, or the report was missing. In addition, some reports did not include the interpretation of the image, as it was a follow-up X-ray. In these cases, the report only indicated whether or not there were changes with respect to the previous report and, therefore, they had to be discarded from the analysis (Fig. 1). Of the final sample, 144 (51.8%) obtained images without radiological abnormalities according to the radiologist's report.

Although the final sample consisted of 278 images, it is possible that an image may be suggestive of one or more conditions or anatomical abnormalities. In this sense, the sum of the unaltered images plus the images with one or more abnormalities does not correspond to the total sample ($n = 278$). Of the conditions or anatomical abnormalities for which the algorithm was trained, the AI model identified 33 in the sample, and the most prevalent were: nodule ($n = 37$, 13.3%), consolidation ($n = 28$, 10.1%), abnormal rib ($n = 17$, 6.1%), enlarged heart ($n = 15$, 5.4%) and aortic sclerosis ($n = 8$, 2.9%). On the other hand, the reference radiologist identified 35 in the sample, and the most prevalent were: consolidation ($n = 30$, 10.8%), pulmonary emphysema ($n = 20$, 7.2%), enlarged aorta ($n = 14$, 5.0%), enlarged heart ($n = 12$, 4.3%), hilar prominence ($n = 12$, 4.3%) and linear atelectasis ($n = 11$, 3.9%) (Table 1).

Analysing the validity of the AI algorithm, the average accuracy was 0.95 (95% CI 0.92; 0.98), the average sensitivity was 0.48 (95% CI 0.30; 0.66) and the average specificity was 0.98 (95% CI 0.97; 0.99). The accuracy, sensitivity and specificity values for each condition can be seen in Table 2. The values for true positives, true negatives, false positives, and false negatives for each condition are presented in Table 1 of the supplementary information.

The reference radiologist identified a list of conditions for which the algorithm was not trained, and which were therefore classified as "Other". The most prevalent were bronchial wall thickening ($n = 13$, 4.68%), fibrosclerotic lesions or abnormalities ($n = 11$, 3.96%) and chronic pulmonary abnormalities ($n = 11$, 3.96%) (Table 3).

Figures 2 and 3 show some examples of the AI algorithm's performance in cases where the algorithm's diagnosis was successful and in cases where errors occurred.

In order to perform a more general analysis, the conditions were grouped into 10 groups according to chest anatomy, considering only the diagnoses for which the AI model was trained. According to the radiologist, the most prevalent groupings found were lung parenchymal conditions ($n = 71$, 25.5%), bone conditions ($n = 21$,

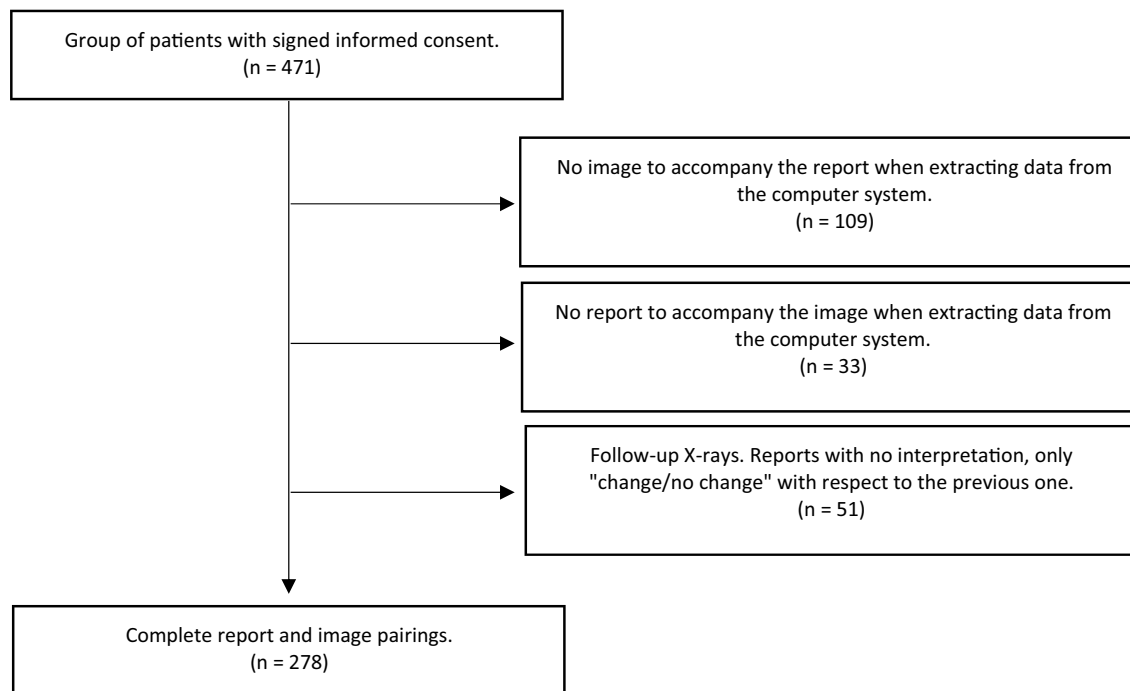


Figure 1. Flow chart of the final sample of study images.

7.5%) and vessel conditions ($n = 18$, 6.47%). According to the AI model, the most prevalent conditions were lung parenchymal conditions ($n = 57$, 20.5%), bone conditions ($n = 21$, 7.5%) and cardiac and/or valvular conditions ($n = 15$, 5.4%) (Table 4).

Finally, Table 5 shows the accuracy, sensitivity and specificity values for each group. Of these, it is worth mentioning the low sensitivity values for mediastinal conditions (0.0, 95% CI 0.0; 0.96), vessel conditions (0.29, 95% CI 0.11; 0.52) and bone conditions (0.24, 95% CI 0.08; 0.47). On the other hand, high sensitivity values were recorded for external implants (0.67, 95% CI 0.22; 0.96), upper abdominal conditions (0.67, 95% CI 0.30; 0.93) and cardiac and/or valvular conditions (0.67, 95% CI 0.35; 0.90). It is also worth mentioning the model's strong ability to detect images that do not have radiological abnormalities. The values for true positives, true negatives, false positives, and false negatives for each grouping are presented in Table 1 of the supplementary information.

Discussion

The aim of this study was to perform an external validation, in real clinical practice, of the diagnostic capability of an AI algorithm with respect to the reference radiologist for chest X-rays, as well as to detect possible diagnoses for which the algorithm had not been trained. Thus, the overall accuracy of the algorithm was 0.95 (95% CI 0.92–0.98), the sensitivity was 0.48 (95% CI 0.30–0.66) and the specificity was 0.98 (95% CI 0.97–0.99). The results obtained have further highlighted, as indicated by different expert groups^{26,28,29}, the need for external validations of AI algorithms in a real clinical context in order to establish the necessary measures and adaptations to ensure safety and effectiveness in any environment. Therefore, in the context of the model developed, it is important to understand and interpret what each of the results obtained indicate.

High accuracy values were observed in most cases (ranging between 0.7–1). The accuracy is represented by the proportion of correctly classified results among the total number of cases examined. This value was high since, both for each condition and for the groups of conditions, the capacity to detect true negatives was good, taking into account that most of the images analysed were found to have no abnormalities (51.8%). Working with an AI algorithm that quickly determines that there is no abnormality can function as a triage tool, streamlining the diagnostic process, allowing the professional to focus on other tests, reduce waiting lists, reduce waiting times for diagnoses and even reduce expenses in secondary tests.

With sensitivity referring to the ability to detect an abnormality when there really is one, high sensitivity values were shown when detecting anatomical findings or abnormalities such as sternal cables, enlarged heart, abnormal ribs, spinal implants, cardiac valve, or interstitial markings. On the other hand, low sensitivity values were observed for most conditions, indicating that the algorithm had limited ability to detect certain conditions like those in the mediastinum, vessels, or bones. These findings align with the results of a study that performed an external validation of a similar algorithm in an emergency department³⁵. Additionally, the algorithm exhibited low sensitivity in detecting pulmonary emphysema, linear atelectasis, and hilar prominence, which are prevalent conditions in the primary care setting³¹.

Low sensitivity was also observed when detecting nodules, with the algorithm finding more nodules than the reference radiologist, in most cases confusing them with areolae in the breast tissue. Although it is important to be able to detect any warning signs and that the professional is in charge of making the clinical judgement and

Condition/finding	Radiologist N (%)	AI algorithm N (%)
Pleural adhesion	–	5 (1.8)
Enlarged aorta	14 (5.04)	1 (0.36)
Linear atelectasis	11 (3.96)	4 (1.44)
Nuss bar or Pectus excavatum	1 (0.36)	–
Sternal wires	3 (1.08)	3 (1.08)
Lymph node calcification	–	2 (0.72)
Spinal degenerative changes	6 (2.16)	–
Enlarged heart	12 (4.32)	15 (5.4)
Kyphosis	3 (1.08)	–
Catheter placement	1 (0.36)	–
Congestion	–	3 (1.08)
Consolidation	30 (10.79)	28 (10.07)
Abnormal rib	3 (1.08)	17 (6.12)
Mediastinal shift	–	1 (0.36)
Hilar prominence	12 (4.32)	–
Elevated diaphragm	4 (1.44)	7 (2.52)
Pulmonary emphysema	20 (7.19)	3 (1.08)
Bullous emphysema	2 (0.72)	–
Pleural thickening	4 (1.44)	–
Fissural thickening	1 (0.36)	2 (0.72)
Spinal enthesopathy	1 (0.36)	–
Aortic sclerosis	2 (0.72)	8 (2.88)
Scoliosis	3 (1.08)	3 (1.08)
Pulmonary fibrosis	1 (0.36)	–
Spinal fracture	6 (2.16)	–
Gastric bubble	–	3 (1.08)
Granuloma	2 (0.72)	4 (1.44)
Hiatal hernia	5 (1.8)	6 (2.16)
Pulmonary hypertension	3 (1.08)	–
Hypoventilation	–	1 (0.36)
Spinal implant	1 (0.36)	1 (0.36)
Lymphadenopathy	2 (0.72)	5 (1.8)
Pacemaker	2 (0.72)	2 (0.72)
Interstitial markings	4 (1.44)	8 (2.88)
Mass	1 (0.36)	–
Widened mediastinum	1 (0.36)	1 (0.36)
Pneumoperitoneum	–	1 (0.36)
Nodule	3 (1.08)	37 (13.31)
Pneumomediastinum	–	3 (1.08)
Pneumothorax	–	1 (0.36)
Sarcoidosis	–	4 (1.44)
Tuberculosis	1 (0.36)	6 (2.16)
Artificial heart valve	1 (0.36)	1 (0.36)
Pleural effusion	10 (3.6)	6 (2.16)
No abnormalities	144 (51.8)	203 (73.02)
Others	63 (22.67)	–

Table 1. Description of the conditions or anatomical abnormalities of the 278 images and their respective diagnoses according to the radiologist and the AI algorithm.

determining the need for complementary tests, it is possible that this external validation has detected a possible gender bias in the training of the algorithm. When it comes to chest imaging, it's important to distinguish between the physiological aspects of breast tissue and any potential changes it may undergo during various life stages, as opposed to signs of conditions or abnormalities³⁶. Other studies have also detected a high false positive value in the detection of nodules due to other causes such as fat, pleura or interstitial lung disease³⁷.

Condition/finding	Accuracy	95% CI	Sensitivity	95% CI	Specificity	95% CI	PPV	95% CI	NPV	95% CI
Average	0.95	(0.92; 0.98)	0.48	(0.30; 0.66)	0.98	(0.97; 0.99)				
Pleural adhesion	–	–	–	–	–	–	–	–	–	–
Enlarged aorta	0.95	(0.92; 0.97)	0.07	(0.00; 0.34)	1.00	(0.99; 1.00)	1.00	(0.02; 1.00)	0.95	(0.92; 0.97)
Linear atelectasis	0.70	(0.94; 0.98)	0.27	(0.06; 0.61)	0.99	(0.98; 1.00)	0.75	(0.19; 0.99)	0.97	(0.94; 0.99)
Nuss bar or Pectus excavatum	–	–	–	–	–	–	–	–	–	–
Sternal wires	1.00	(0.98; 1.00)	1.00	(0.29; 1.00)	1.00	(0.99; 1.00)	1.00	(0.29; 1.00)	1.00	(0.99; 1.00)
Lymph node calcification	–	–	–	–	–	–	–	–	–	–
Spinal degenerative changes	–	–	–	–	–	–	–	–	–	–
Enlarged heart	0.96	(0.93; 0.98)	0.67	(0.35; 0.90)	0.97	(0.95; 0.99)	0.53	(0.27; 0.79)	0.98	(0.96; 1.00)
Kyphosis	–	–	–	–	–	–	–	–	–	–
Catheter placement	–	–	–	–	–	–	–	–	–	–
Congestion	–	–	–	–	–	–	–	–	–	–
Consolidation	0.89	(0.86; 0.93)	0.50	(0.31; 0.69)	0.95	(0.91; 0.97)	0.54	(0.34; 0.72)	0.94	(0.90; 0.97)
Abnormal rib	0.94	(0.91; 0.97)	0.67	(0.09; 0.99)	0.95	(0.91; 0.97)	0.12	(0.01; 0.36)	0.99	(0.98; 1.00)
Mediastinal shift	–	–	–	–	–	–	–	–	–	–
Hilar prominence	–	–	–	–	–	–	–	–	–	–
Elevated diaphragm	0.97	(0.94; 0.98)	0.25	(0.01; 0.81)	0.98	(0.95; 0.99)	0.14	(0.00; 0.58)	0.99	(0.97; 1.00)
Pulmonary emphysema	0.93	(0.89; 0.96)	0.10	(0.01; 0.32)	0.99	(0.98; 1.00)	0.67	(0.09; 0.99)	0.93	(0.90; 0.96)
Bullous emphysema	–	–	–	–	–	–	–	–	–	–
Pleural thickening	–	–	–	–	–	–	–	–	–	–
Fissural thickening	0.99	(0.98; 0.99)	1.00	(0.03; 1.00)	0.99	(0.98; 1.00)	0.50	(0.01; 0.99)	1.00	(0.99; 1.00)
Spinal enthesopathy	–	–	–	–	–	–	–	–	–	–
Aortic sclerosis	0.96	(0.93; 0.98)	0.00	(0.00; 0.37)	0.97	(0.94; 0.99)	0.00	(0.00; 0.37)	0.99	(0.97; 1.00)
Scoliosis	0.98	(0.96; 0.99)	0.33	(0.01; 0.91)	0.99	(0.97; 1.00)	0.33	(0.01; 0.91)	0.99	(0.97; 1.00)
Pulmonary fibrosis	–	–	–	–	–	–	–	–	–	–
Spinal fracture	–	–	–	–	–	–	–	–	–	–
Gastric bubble	–	–	–	–	–	–	–	–	–	–
Granuloma	0.98	(0.95; 0.99)	0.00	(0.00; 0.84)	0.98	(0.96; 1.00)	0.00	(0.00; 0.60)	0.99	(0.97; 1.00)
Hiatal hernia	0.99	(0.98; 0.99)	1.00	(0.48; 1.00)	0.99	(0.98; 1.00)	0.83	(0.36; 1.00)	1.00	(0.99; 1.00)
Pulmonary hypertension	–	–	–	–	–	–	–	–	–	–
Hypoventilation	–	–	–	–	–	–	–	–	–	–
Spinal implant	1.00	(0.98; 1.00)	1.00	(0.03; 1.00)	1.00	(0.99; 1.00)	1.00	(0.03; 1.00)	1.00	(0.99; 1.00)
Lymphadenopathy	0.97	(0.95; 0.99)	0.50	(0.00; 0.97)	0.98	(0.96; 0.99)	0.20	(0.00; 0.52)	0.99	(0.98; 1.00)
Pacemaker	0.99	(0.97; 0.99)	0.50	(0.01; 0.99)	0.99	(0.98; 1.00)	0.50	(0.01; 0.99)	0.99	(0.98; 1.00)
Interstitial markings	0.98	(0.95; 0.99)	0.75	(0.19; 0.99)	0.98	(0.96; 0.99)	0.38	(0.09; 0.76)	0.99	(0.98; 1.00)
Mass	–	–	–	–	–	–	–	–	–	–
Widened mediastinum	0.99	(0.97; 0.99)	0.00	(0.00; 0.97)	0.99	(0.98; 1.00)	0.00	(0.00; 0.97)	0.99	(0.98; 1.00)
Pneumoperitoneum	–	–	–	–	–	–	–	–	–	–
Nodule	0.86	(0.82; 0.91)	0.00	(0.00; 0.71)	0.88	(0.84; 0.92)	0.00	(0.00; 0.11)	0.99	(0.96; 1.00)
Pneumomediastinum	–	–	–	–	–	–	–	–	–	–
Pneumothorax	–	–	–	–	–	–	–	–	–	–
Sarcoidosis	–	–	–	–	–	–	–	–	–	–
Tuberculosis	0.98	(0.96; 0.99)	1.00	(0.03; 1.00)	0.98	(0.96; 0.99)	0.17	(0.00; 0.64)	1.00	(0.99; 1.00)
Artificial heart valve	1.00	(0.98; 1.00)	1.00	(0.03; 1.00)	1.00	(0.99; 1.00)	1.00	(0.02; 1.00)	1.00	(0.99; 1.00)
Pleural effusion	0.97	(0.94; 0.98)	0.40	(0.12; 0.74)	0.99	(0.97; 1.00)	0.67	(0.22; 0.96)	0.98	(0.95; 0.99)
Others	–	–	–	–	–	–	–	–	–	–

Table 2. Accuracy, sensitivity, specificity, positive and negative predictive values with a 95% confidence interval for each condition. *PPV* positive predictive value; *NPV* negative predictive value.

Finally, specificity being the ability to correctly identify images in which there are no radiological abnormalities, the results showed high values for all condition groupings, since the algorithm was able to detect images with no abnormalities.

Following the authors' desire to contribute to the improvement of the AI model, some radiologists' findings were identified that were overlooked during the algorithm's training, especially related to bronchial conditions, including chronic bronchopathy, bronchiectasis, and bronchial wall thickening. Additionally, the algorithm missed common chronic conditions often seen in primary care, including chronic pulmonary abnormalities,

Condition/finding	N (%)
Air bronchogram	1 (0.36)
Chronic bronchopathy/COPD	4 (1.44)
Bronchiectasis	6 (2.16)
Aortic calcification	1 (0.36)
Post-surgical abnormalities	4 (1.44)
Chronic pulmonary abnormalities	11 (3.96)
Chondrocalcinosis	1 (0.36)
Bronchial wall thickening	13 (4.68)
Ankylosing spondylitis	1 (0.36)
Nonspecific	2 (0.72)
Fibrocystic lesions or fibrocystic abnormalities	11 (3.96)
Diaphragmatic eventration	6 (2.16)
Breast prostheses	1 (0.36)
Linear opacities	1 (0.36)

Table 3. Description of conditions not contemplated by the AI model.

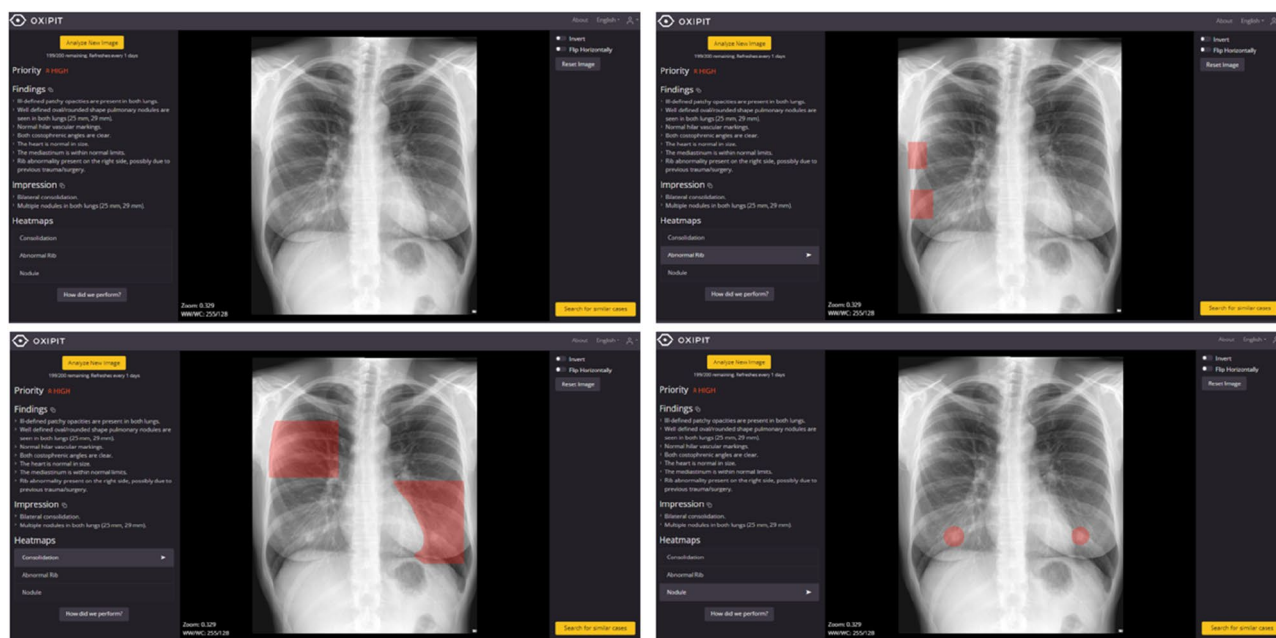


Figure 2. Image of patient (upper-left) where according to the radiologist’s report there is only consolidation, but the algorithm detects an abnormal rib (upper-right), consolidation (lower-left) and two nodules (lower-right). It is worth noting the confusion of a consolidation with mammary tissue and of two nodules with the two mammary areolae.

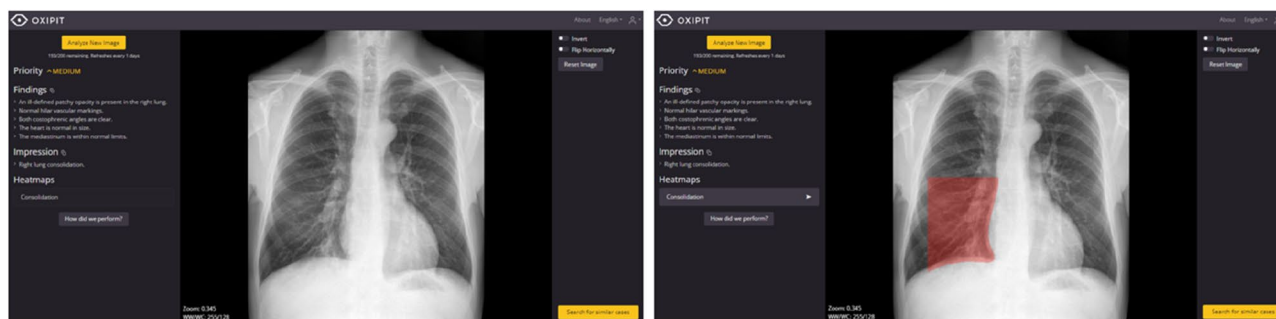


Figure 3. Image of patient (left) where the AI algorithm and the radiologist detected the same condition: consolidation (right).

Grouping	Radiologist N (%)	AI algorithm N (%)
Others	3 (1.08)	9 (3.24)
External implants	6 (2.16)	5 (1.80)
Mediastinum	1 (0.36)	4 (1.44)
Upper abdomen conditions	9 (3.24)	13 (4.68)
Cardiac and/or valvular conditions	12 (4.32)	15 (5.40)
Vessel conditions	18 (6.47)	9 (3.24)
Bone conditions	21 (7.55)	21 (7.55)
Pulmonary parenchymal conditions	71 (25.50)	57 (20.5)
Pleural conditions	13 (4.68)	12 (4.32)
No abnormalities	144 (51.8)	203 (73.0)

Table 4. Description of the conditions of the 278 images according to the radiologist and AI model, grouped according to chest anatomy.

Condition/finding	Accuracy	95% CI	Sensitivity	95% CI	Specificity	95% CI
Others	0.96	(0.93; 0.98)	0.33	(0.01; 0.91)	0.97	(0.94; 0.99)
External implants	0.99	(0.97; 0.99)	0.67	(0.22; 0.96)	0.99	(0.98; 1.00)
Mediastinum	0.98	(0.96; 0.99)	0.00	(0.00; 0.97)	0.99	(0.96; 1.00)
Upper abdomen conditions	0.96	(0.93; 0.98)	0.67	(0.30; 0.93)	0.97	(0.95; 0.99)
Cardiac and/or valvular conditions	0.96	(0.93; 0.98)	0.67	(0.35; 0.90)	0.97	(0.95; 0.99)
Vessel conditions	0.92	(0.89; 0.95)	0.29	(0.11; 0.52)	0.99	(0.97; 1.00)
Bone conditions	0.89	(0.84; 0.92)	0.24	(0.08; 0.47)	0.94	(0.90; 0.96)
Pulmonary parenchymal conditions	0.78	(0.72; 0.82)	0.46	(0.35; 0.59)	0.88	(0.83; 0.92)
Pleural conditions	0.95	(0.92; 0.97)	0.46	(0.19; 0.75)	0.98	(0.95; 0.99)
With abnormalities VS without	0.70	(0.64; 0.75)	0.47	(0.38; 0.56)	0.92	(0.86; 0.96)

Table 5. Accuracy, sensitivity and specificity values for each grouping.

COPD, and fibrocystic abnormalities. Furthermore, it was noted that certain condition names within the AI algorithm should be adjusted to align with names used in the radiology field. Interstitial markings could be changed to interstitial abnormality, consolidation to condensation, aortic sclerosis to valvular sclerosis, and abnormal rib to rib fracture.

Once the main variables that characterise the algorithm's capacity were discussed, the results obtained differ from the majority of published studies, since most of them obtained a higher algorithm capacity. However, it should be noted that most of these are internal validations and not tested in real clinical practice settings^{38–40}.

A study in Korea performed an internal and external validation of an AI algorithm capable of detecting the 10 most prevalent chest X-ray abnormalities and was able to demonstrate the difference in sensitivity and specificity values. The internal validation obtained sensitivity and specificity values between 0.87–0.94 and 0.81–0.98, respectively. On the other hand, the external validation obtained sensitivity and specificity values between 0.61–1.00 and 0.71–0.98, respectively⁴¹. This difference can also be seen in a study in Michigan, where internal and external validation of an AI algorithm capable of detecting the most common chest X-ray abnormalities was performed⁴², and in a study at the Seoul University School of Medicine, where an algorithm for lung cancer detection in population screening was validated⁴³.

Therefore, the results obtained from the external validation show the need to increase the sensitivity of the algorithm for most conditions. Considering that AI should serve as a diagnostic support tool and the ultimate responsibility for medical decisions rests with the practitioner, it is ideal for the algorithm to flag potential abnormalities for the practitioner to review and confirm. This ensures the highest diagnostic effectiveness. Recent studies have shown that the use of an AI algorithm to support the practitioner significantly improves diagnostic sensitivity and specificity and reduces image reading time^{20,44}.

Enhanced sensitivity could help address the shortage of specialised radiologists globally, especially in Central Catalonia's primary care setting, where this validation was conducted^{45,46}. More and more, general practitioners are tasked with interpreting X-rays. In this context, the advancement of these tools can be a valuable asset in the diagnostic process.

Limitations and strengths

One significant limitation of the study was the small sample size for certain specific conditions. This was due to difficulties in obtaining the required number of cases, as these conditions are not very common in real clinical practice. Consequently, the external validation for these conditions yielded less reliable estimates. However, by

representing reality, a large volume of images without radiological abnormalities was obtained and this allowed for a good external validation of the model's ability to detect images without abnormalities.

In addition, the radiologist's reference diagnosis was not always the practitioner's own, but that of a group of radiologists. This could represent a limitation, since there was no consensus among them, but there was no desire to alter actual clinical practice. In addition, the study aimed to test the algorithm in primary care settings. For this reason, a double interpretation of the images was performed: initially by the radiologist and subsequently, the radiologist's report was interpreted by the family and community physician. Finally, another limitation was the lack of information on the sex of the users analysed. Through the results obtained, we found it very relevant to do another study but separating the capabilities of the algorithm according to gender, since it seems that they might not be the same. In addition, since we have a small sample for most of the conditions, separating the analyses according to sex in the present study would be unreliable and not comparable.

On the other hand, the greatest strength of the study is that it presents an external validation in real clinical practice in primary care and there are currently few studies that have done so. Most studies present an internal validation, but it is very important to perform an external validation in order to estimate the accuracy of the model in a population other than the training population, thus allowing the results to be generalised.

Conclusion

The findings of this study demonstrate the validation of an AI algorithm for reading chest X-rays in the primary care setting, achieved by comparing its diagnoses with those made by a radiologist. The algorithm has been validated in the primary care setting using values such as the accuracy, sensitivity and specificity of the algorithm and has proven to be useful by being able to identify images with or without abnormalities. However, further training is needed to increase the diagnostic capability of some of the conditions analysed. It is important that training is done in a real environment, with real images, in order to perform robust external validations. Our analysis highlights the need for continuous improvement to ensure that the algorithm is a reliable and effective tool in the primary care environment.

The role of AI in healthcare should be to assist and support the practitioners. Being able to reliably detect images without abnormalities can have a very positive impact, reducing waiting times for diagnoses, secondary tests to rule out conditions, streamlining practitioners work and, among others, ultimately favouring patient care and, indirectly, their health.

Data availability

The datasets generated and/or analysed during the current study are not publicly available because our manuscript was based on confidential and sensitive health data. However, they are available from the corresponding author upon reasonable request.

Received: 30 October 2023; Accepted: 27 February 2024

Published online: 03 March 2024

References

- Novelline, R. A. & Squire, L. F. Squire's fundamentals of radiology. La Editorial U, editor. (2004).
- W H. Learning radiology: Recognizing the basics. Sciences EH, editor. (2015).
- Hwang, E. J. *et al.* Development and validation of a deep learning-based automated detection algorithm for major thoracic diseases on chest radiographs. *JAMA Netw. Open* **2**(3), e191095. <https://doi.org/10.1001/jamanetworkopen.2019.1095> (2019).
- Wu, J. T. *et al.* Comparison of chest radiograph interpretations by artificial intelligence algorithm vs radiology residents. *JAMA Netw. Open* **3**(10), e2022779. <https://doi.org/10.1001/jamanetworkopen.2020.22779> (2020).
- Santos, Á. & Solís, P. Posición SERAM sobre la necesidad de informar la radiología simple. *Soc Española Radiol Médica* **1**, 1–47 (2015).
- UNSCEAR. Sources and effects of ionizing radiation volume I: Source. Vol. I, United Nations Scientific Committee on the Effects of Atomic Radiation. 1–654 (2000).
- Rimmer, A. Radiologist shortage leaves patient care at risk, warns royal college. *BMJ* <https://doi.org/10.1136/bmj.j4683> (2017).
- Bhargavan, M., Sunshine, J. H. & Schepps, B. Too few radiologists?. *Am. J. Roentgenol.* **178**(5), 1075–1082 (2002).
- Chew, C., O'Dwyer, P. J. & Young, D. Radiology and the medical student: Do increased hours of teaching translate to more radiologists?. *BJR Open* **3**(1), 20210074 (2021).
- Lyon, M. *et al.* Rural ED transfers due to lack of radiology services. *Am. J. Emerg. Med.* **33**(11), 1630–1634 (2015).
- European A. Radiology services in Europe : Harnessing growth is health system dependent (2022).
- Patlas, M. N., Katz, D. S. & Scaglione, M. Errors in Emergency and Trauma Radiology. (Springer, 2007). Available from: [https://download.bibis.ir/Books/Medical/_old/Errors in Emergency and Trauma Radiology.pdf](https://download.bibis.ir/Books/Medical/_old/Errors%20in%20Emergency%20and%20Trauma%20Radiology.pdf). Accessed 28 Sep 2023.
- Ruutinen, A. T., Durand, D. J., Scanlon, M. H. & Itri, J. N. Increased error rates in preliminary reports issued by radiology residents working more than 10 consecutive hours overnight. *Acad. Radiol.* **20**(3), 305–311 (2013).
- Hanna, T. N. *et al.* Emergency radiology practice patterns: Shifts, schedules, and job satisfaction. *J. Am. Coll. Radiol.* **14**(3), 345–352 (2017).
- Bruls, R. J. M. & Kwee, R. M. Workload for radiologists during on-call hours: Dramatic increase in the past 15 years. *Insights Imaging* **11**(1), 1–7 (2020).
- Sociedad Española de Medicina de Familia y Comunitaria (semFYC). Programa de la especialidad de medicina familiar y comunitaria [Internet]. Ministerio de Sanidad y Consumo. 2005. Available from: <https://www.semfyec.es/wp-content/uploads/2016/09/Programa-Especialidad.pdf>. Accessed 28 Sep 2023.
- Koenigkam Santos, M. *et al.* Artificial intelligence, machine learning, computer-aided diagnosis, and radiomics: Advances in imaging towards to precision medicine. *Radiol. Bras.* **52**(6), 387–396 (2019).
- Neri, E. *et al.* What the radiologist should know about artificial intelligence – an ESR white paper. *Insights Imaging* **10**(1), 44 (2019).
- Hwang, E. J. *et al.* Use of artificial intelligence-based software as medical devices for chest radiography: A position paper from the Korean society of thoracic radiology. *Korean J. Radiol.* **22**(11), 1743–1748 (2021).

20. Kim, J. H., Han, S. G., Cho, A., Shin, H. J. & Baek, S. E. Effect of deep learning-based assistive technology use on chest radiograph interpretation by emergency department physicians: a prospective interventional simulation-based study. *BMC Med. Inform. Decis. Mak.* **21**(1), 1–9 (2021).
21. Kaviani, P. *et al.* Performance of a chest radiography AI algorithm for detection of missed or mislabeled findings: A multicenter study. *Diagnostics* **12**(9), 2086 (2022).
22. Bleeker, S. E. *et al.* External validation is necessary in prediction research: A clinical example. *J. Clin. Epidemiol.* **56**(9), 826–832 (2003).
23. Kim, D. W., Jang, H. Y., Kim, K. W., Shin, Y. & Park, S. H. Design characteristics of studies reporting the performance of artificial intelligence algorithms for diagnostic analysis of medical images: Results from recently published papers. *Korean J. Radiol.* **20**(3), 405–410 (2019).
24. Dupuis, M., Delbos, L., Veil, R. & Adamsbaum, C. External validation of a commercially available deep learning algorithm for fracture detection in children. *Diagn. Interv. Imaging* **103**(3), 151–159 (2022).
25. Mutasa, S., Sun, S. & Ha, R. Understanding artificial intelligence based radiology studies: What is overfitting?. *Clin. Imaging* **65**, 96–99 (2020).
26. Liu, X. *et al.* Reporting guidelines for clinical trial reports for interventions involving artificial intelligence: The CONSORT-AI extension. *Nat. Med.* **2**(10), e537–e548 (2020).
27. Kim, C. *et al.* Multicentre external validation of a commercial artificial intelligence software to analyse chest radiographs in health screening environments with low disease prevalence. *Eur. Radiol.* **33**(5), 3501–3509 (2023).
28. Reglamento (UE) 2017/745 DEL PARLAMENTO EUROPEO Y DEL CONSEJO-de 5 de abril de 2017-sobre los productos sanitarios, por el que se modifican la Directiva 2001/83/CE, el Reglamento (CE) n.o 178/2002 y el Reglamento (CE) n.o 1223/2009 y por el que se derogan las Directivas 90/385/CEE y 93/42/CEE del Consejo.
29. Vasey, B. *et al.* Reporting guideline for the early-stage clinical evaluation of decision support systems driven by artificial intelligence: DECIDE-AI. *Nat. Med.* **28**(5), 924–933 (2022).
30. Taylor, M. *et al.* Raising the bar for randomized trials involving artificial intelligence: The SPIRIT-artificial intelligence and CONSORT-artificial intelligence guidelines. *J. Invest. Dermatol.* **141**(9), 2109–2111 (2021).
31. van Beek, E. J. R., Ahn, J. S., Kim, M. J. & Murchison, J. T. Validation study of machine-learning chest radiograph software in primary and emergency medicine. *Clin. Radiol.* **78**(1), 1–7. <https://doi.org/10.1016/j.crad.2022.08.129> (2023).
32. Catalina, Q. M., Fuster-Casanovas, A., Solé-Casals, J. & Vidal-Alaball, J. Developing an artificial intelligence model for reading chest X-rays protocol for a prospective validation study. *JMIR Res. Protoc.* **11**(11), e39536 (2022).
33. Oxipit. Oxipit ChestEye secures medical device certification in Australia. 2020. Available from: <https://oxipit.ai/news/oxipit-ai-medical-imaging-australia/>. Accessed 22 Sep 2023.
34. Oxipit. Study: AI found to reduce bias in Radiology Reports. 2019. Available from: <https://oxipit.ai/news/study-ai-found-to-reduce-bias-in-radiology-reports/>. Accessed 22 Sep 2023.
35. Hwang, E. J. *et al.* Deep learning for chest radiograph diagnosis in the emergency department. *Radiology* **293**(3), 573–580. <https://doi.org/10.1148/radiol.2019191225> (2019).
36. Salas Pérez, R., Teixidó Vives, M., Picas Cutrina, E. & Romero, N. I. Diferentes aspectos de las calcificaciones mamarias. *Imagen Diagnóstica* **4**(2), 52–57 (2013).
37. Koo, Y. H. *et al.* Extravalidation and reproducibility results of a commercial deep learning-based automatic detection algorithm for pulmonary nodules on chest radiographs at tertiary hospital. *J. Med. Imaging Radiat. Oncol.* **65**(1), 15–22 (2021).
38. Futoma, J., Simons, M., Panch, T., Doshi-Velez, F. & Celi, L. A. The myth of generalisability in clinical research and machine learning in health care. *Lancet Digit. Heal.* **2**(9), e489–e492 (2020).
39. Nam, J. G. *et al.* Development and validation of deep learning-based automatic detection algorithm for malignant pulmonary nodules on chest radiographs. *Radiology* **290**(1), 218–228 (2019).
40. Park, S. H. Diagnostic case-control versus diagnostic cohort studies for clinical validation of artificial intelligence algorithm performance. *Radiology* **290**, 272–273 (2019).
41. Nam, J. G. *et al.* Development and validation of a deep learning algorithm detecting 10 common abnormalities on chest radiographs. *Eur. Respir. J.* **57**(5), 2003061 (2021).
42. Sjoding, M. W. *et al.* Deep learning to detect acute respiratory distress syndrome on chest radiographs: A retrospective study with external validation. *Lancet Digit. Heal.* **3**(6), e340–e348 (2021).
43. Lee, J. H. *et al.* Performance of a deep learning algorithm compared with radiologic interpretation for lung cancer detection on chest radiographs in a health screening population. *Radiology* **297**(3), 687–696. <https://doi.org/10.1148/radiol.2020201240> (2020).
44. Sung, J. *et al.* Added value of deep learning-based detection system for multiple major findings on chest radiographs: A randomized crossover study. *Radiology* **299**(2), 450–459 (2021).
45. Sequía de radiólogos en España con plantillas al 50%. Available from: <https://www.redaccionmedica.com/secciones/radiologia/marti-de-gracia-vivimos-una-situacion-critica-de-escasez-de-radiologos--4663>. Accessed 25 Jun 2022.
46. Esquerrà, M. *et al.* Abdominal ultrasound: A diagnostic tool within the reach of general practitioners. *Aten primaria/Soc Española Med Fam y Comunitaria* **44**(10), 576–583 (2012).

Acknowledgements

The authors would like to thank all the users who agreed to participate in the study, and the radiology team at the Osona study centre for their help in recruiting patients. We would also like to thank the general practitioners for their participation in the interpretation of the radiologist's reports. Finally, we would also like to thank the professionals at Oxipit for their help in describing the more technical part of the algorithm. In addition, this study was carried out as part of the Industrial Doctorates programme of Catalonia and obtained a Bayès Grant.

Author contributions

Q.M.C., A.F.C., J.V.A. and J.S.C. participated in the definition of the project idea and the realisation of the protocol. Q.M.C. and A.F.C. participated in data preparation and Q.M.C. performed the data analysis and interpretation. J.V.A., A.R.C. and A.E.B. performed the interpretation of the radiologist's reports for all X-rays in the study. A.F.C., J.V.A., J.S.C., A.R.C. and A.E.B. participated in the drafting of the article and extensive review.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1038/s41598-024-55792-1>.

Correspondence and requests for materials should be addressed to J.V.-A. or J.S.-C.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2024

6. DISCUSSIÓ

Els resultats de la present tesi doctoral descriuen el coneixement i percepció dels usuaris i professionals d'atenció primària sobre l'ús de la IA com a eina de suport assistencial, especialment en l'àmbit de la radiologia, i mostren la validació externa d'una eina d'IA per a la lectura de plaques de tòrax en l'atenció primària.

6.1 Coneixements i percepció de la població envers l'ús de la IA en la salut general i en el camp de la radiologia

Gairebé nou de cada deu ciutadans indica conèixer el concepte d'IA. Avui en dia, aquest resultat no resulta estrany ja que la IA cada vegada es troba més present en el dia a dia de la població. Tot i això, altres estudis on s'ha demanat sobre coneixements més específics observen un desconeixement generalitzat [25,40,140,141]. Una possible explicació a aquesta discrepància és que el coneixement que indica tenir la ciutadania no és específic, sinó simples nocions obtingudes sobretot d'haver-ne sentit a parlar i no pas d'experiències personals o de formació [142].

Així doncs, i tot i que és d'esperar que cada vegada la ciutadania conegui més a fons la IA degut a la seva proximitat, popularitat i atenció als mitjans de comunicació, és important posar l'accent en la correcta formació de les persones sobre la IA, ja que diversos estudis han demostrat que el desconeixement sovint és motiu de rebuig i desconfiança. A més, aquests mateixos estudis han mostrat que la població indica en diverses ocasions una necessitat de formació en temes com el raonament i lògica que utilitza la IA, les seves fonts d'informació, la fiabilitat dels seus algorismes i la despersonalització. Per tant, és de vital importància crear estratègies per fer arribar un bon coneixement a tota la població [21,39,140,142–146].

En aquest sentit, hi ha un debat global sobre quina informació s'ha de donar, com i a qui. Alguns defensen que és important començar a treballar-hi des de la infància per poder combatre el problema des de l'arrel i en canvi, d'altres opinen que la formació hauria de ser

més específica per a cada àmbit i, en el cas de la salut, caldria donar-la als propis graus universitaris o bé com a formacions extres d'actualització als professionals [39].

D'altra banda, i tot i les discrepàncies en el coneixement de la IA, els resultats han mostrat un sentiment lleugerament negatiu de l'aplicació de la IA en l'àmbit de la salut i la radiologia per la possible pèrdua de la interacció personal amb el professional, per la necessitat d'estar informat, per la desconfiança envers la IA i per la falta de responsabilitat d'aquesta. Una possible explicació a aquesta baixa puntuació pot ser la falta de coneixement específic sobre la IA, perquè s'ha observat que la població amb nivells més alts d'educació i la població amb coneixements previs sobre la IA presenta menys preocupació en aquestes àrees, coincidint aquests resultats amb altres estudis [28,29,39,146,147].

En efecte, resultats d'altres estudis concorden amb els nostres observant certa reticència en l'aplicació de la IA, coincidint les àrees amb més preocupació les quals son la pèrdua de la interacció social, la falta de la supervisió humana, la possibilitat que el professional sanitari no dediqui el temps necessari en el cas, la despersonalització i la pèrdua de la protecció de les dades. Així mateix, la ciutadania confia més en els professionals que en la IA i veu molt clar que les aplicacions d'IA en cap cas han de substituir el professional ni prendre responsabilitats, sinó que han de ser un suport i una ajuda a la decisió final del professional [23,135,142,145–148].

En particular, els nivells de preocupació respecte la desconfiança amb la IA i la seva falta de responsabilitat han sigut diferents segons el sexe i l'edat de la població, sent el sexe femení i la població major de 65 anys qui mostra percepcions més negatives. Aquests resultats concorden amb altres estudis on també s'ha observat major desconfiança en la IA en població d'edat avançada i en persones de sexe femení [40,145,148].

En relació a l'edat, caldria desenvolupar estudis per aprofundir més en l'anàlisi de les diferències observades, ja que és possible que no només es tracti de l'edat en sí, sinó que hi hagi amagades variables de confusió com la naturalesa dels problemes de salut en els diferents grups d'edat, la necessitat de tenir contacte personal en la població més envellida

degut a la soledat, la cronicitat o multimorbiditat, o fins i tot el pes de l'experiència i la costum de les persones d'edat més avançada [148,149].

A més, la població més envellida també és, sovint, la més afectada per l'anomenada esclletxa digital, les dificultats en l'accés i el domini del món digital. En aquest sentit, des de la UE ja s'ha posat en marxa un programa anomenat "Configurar el futur digital d'Europa" amb l'objectiu de poder abordar aquestes diferències i fomentar la inclusió per garantir que tothom pugui contribuir al món digital i beneficiar-se d'ell [150].

En relació al sexe també caldria seguir investigant el motiu de les diferències ja que segurament, i com amb l'edat, hi ha altres variables de confusió que poden ajudar en la seva explicació. Un possible motiu podria ser que els homes fan menys ús del sistema sanitari i per tant, no perceben l'afectació que podria tenir l'aplicació de la IA com les dones tal i com s'ha observat a diferents estudis [148].

Si bé les percepcions de la ciutadania sobre la utilització de la IA en salut i especialment en l'àmbit de la radiologia han sigut lleugerament negatives, és també important destacar les àrees on s'ha indicat certa esperança i positivitat respecte la IA. En general, s'ha observat un sentiment d'esperança en relació a l'eficiència que pot oferir la IA, coincidint amb altres estudis on s'ha indicat una percepció positiva en relació a la possible reducció dels temps d'espera per a la realització de proves, el temps d'espera per obtenció del resultats, l'ajuda en la prevenció d'errors mèdics, el suport en la planificació del tractament basat en l'evidència científica, l'àmplia disponibilitat, la facilitat d'ús i la capacitat de reduir el cost de la prestació dels serveis sanitaris [39,40,142,145–148].

A més, tot i no ser un resultat directe de la present tesi, resulta interessant destacar que a altres estudis s'ha observat que la ciutadania creu que la IA pot estar més capacitada que els professionals sanitaris per mantenir-se actualitzada [147,151]. Aquest fet resulta interessant ja que l'àmbit de la salut és un camp en constant evolució, i això implica que els professionals s'han de mantenir en constant formació per no quedar desactualitzats, tant de conceptes teòrics com pràctics. Així doncs, la població considera que la IA podria ser una eina de suport en aquesta constant actualització.

Una bona mostra del sentiment positiu respecte l'ús de la IA tot i els resultats de la present tesi n'és una *scoping review* de l'any 2022 on es conclou, després d'analitzar vint-i-sis articles, que la majoria mostren una acceptabilitat almenys moderada de l'ús de la IA i una acceptació majoritària en problemes de salut menors. Així mateix i d'acord amb els resultats obtinguts, també va concloure que l'edat, els coneixements sobre la IA, l'experiència prèvia i la confiança percebuda son factors que influeixen en l'acceptabilitat de la IA. De fet, ja s'ha demostrat que les persones amb una visió positiva de la IA i una percepció d'eficiència respecte l'ús d'aquesta tenen major confiança en la seva aplicació [39,147].

També, un estudi recent del Regne Unit va fer un prova pilot a diferents centres d'atenció primària sobre l'ús d'una aplicació de mòbil anomenada Ada, creada a partir d'IA, que tenia per objectiu ajudar a l'usuari a saber quin tipus d'atenció i urgència requeria el seu dolor, és a dir, fer de triatge. A part d'avaluar el funcionament de l'aplicació, es va observar que un cop utilitzada l'aplicació, gairebé tots els usuaris trobaven fàcil o molt fàcil el seu ús, la tornarien a utilitzar i la recomanarien a amics o familiars, especialment entre la població més jove [149].

És a dir, tot i que la ciutadania a priori pot tenir una visió negativa sobre la aplicació de la IA en l'àmbit de la salut, el desconeixement podria ser un dels principals motius d'aquesta visió, ja que el coneixement descrit sobre aquesta és molt poc profund i en la majoria dels casos, no està basat en la pròpia experiència del seu ús sinó en prejudicis, preconcepcions i suggestions. D'aquest raonament en son un bon exemple els resultats de la present tesi, on les persones amb coneixements previs sobre la IA i amb majors nivells de formació mostren major capacitat per veure els aspectes positius de l'aplicació de la IA, major acceptació d'aquesta i menor sentiment de reticència.

Precisament, una enquesta sobre l'ús de la IA durant la pandèmia de la Covid-19 realitzada a la població jove d'Àrab Saudita amb un ús diari d'aplicacions d'IA, va mostrar percepcions positives d'aquesta, evidenciant la importància de disposar d'eines de suport als professionals durant pandèmies o situacions amb molt volum de pacients. Aquests resultats reforcen que tant l'edat com la familiarització amb la IA son factors que influeixen a tenir visions més positives envers l'ús de la IA [152].

Per tant, oferir i dotar a la població d'un bon coneixement sobre la IA i d'experiències d'ús real al seu dia a dia amb aplicacions d'IA podria revertir l'actual rebuig i desconfiança. Com indica l'estudi d'Àrabia Saudita, quan l'usuari té una experiència positiva amb una aplicació de IA, la percepció respecte aquesta es veu afectada positivament, especialment en la població més jove. Aquest resultat també es conclou a un estudi realitzat pel New York Times on es va indicar que els debats dels últims trenta anys sobre la IA sempre havien acabat sent positius [153], i seguint amb aquests resultats, alguns estudis també han indicat que la població té percepcions negatives d'aplicacions de la IA fins que està en contacte amb una de veritat i la utilitza. A partir del seu ús, la seva visió canvia i l'experiència passa a ser, en la majoria dels casos, positiva [40,142].

En conseqüència, tot i que a dia d'avui encara hi ha poca literatura on s'estudien les percepcions i els coneixements de la societat sobre la implementació de la IA, cal seguir desenvolupament estudis per crear estratègies i poder aproximar la IA a la població millorant-ne la seva percepció, la seva confiança, el seu coneixement sobre aquesta així com brindant la oportunitat de poder utilitzar-la al dia a dia en primera persona [21,22].

Segons Everett M. Rogers, per tal de fer arribar la informació de qualsevol innovació tecnològica a tota la població cal fer una bona difusió a través de diferents canals de comunicació i durant un període prolongat en el temps per tal de fer un correcte abordatge als cinc passos definits en el procés de la difusió d'innovació: el coneixement, la persuasió, la decisió, la implementació i la confirmació [154].

Inevitablement, la IA seguirà evolucionant en els pròxims anys sent més present en el dia a dia de la ciutadania, i segurament, la ciutadania també tindrà una visió cada vegada més positiva i un coneixement més avançat. El més important, i a la vegada complicat, serà que aquests dos creixements vagin a la mateixa velocitat.

6.2 Coneixements i percepció dels professionals sanitaris envers l'ús de la IA en la salut

En relació als professionals sanitaris, hi ha un gran volum d'ells que indica conèixer el concepte de la IA. Gairebé nou de cada deu entén el concepte de la IA, però no hi ha un coneixement generalitzat sobre el seu ús, perquè de cada deu professionals, quatre consideren que l'utilitzen al seu dia a dia laboral, tres consideren que no l'utilitzen i dos no ho saben. És a dir, tot i l'elevat volum de professionals que indica conèixer el que significa la IA, aquesta ambigüitat sobre el seu ús demostra una clara falta de coneixement profund, formació específica sobre la IA i transparència dels programes sanitaris utilitzats a la consulta.

De fet, en contra dels resultats de la present tesi, altres estudis han indicat que el percentatge de professionals amb el coneixement necessari sobre la IA és baix i en certs casos, massa centrat en coneixements teòrics i no pràctics degut a la falta d'experiències reals amb aquestes eines [46,47,155–158]. A més, el percentatge d'ús de la IA estimat no s'ajusta al de la literatura, on s'estima que el percentatge de professionals que fan actualment ús de la IA al seu dia a dia de la consulta es troba entre el 10% i el 30% [43,46,157–161].

Una possible explicació d'aquest ampli percentatge de professionals que assegura tenir-ne coneixement i de l'ambigüitat del seu ús pot ser la particularitat del territori i la seva aproximació a la salut digital, ja que a la Catalunya Central existeix una llarga tradició d'innovació en *ehealth* essent una de les pioneres en l'ús i aplicació de la telemedicina i tenint gran part del territori en zones rurals allunyades dels hospitals de referència on és primordial disposar d'eines de suport [162]. Recentment, a banda d'aquest estudi, s'ha desenvolupat dos estudis al territori sobre la validació d'algorismes d'IA per la lectura d'imatges de retinopatia diabètica i de dermatologia, i això pot haver sigut una font de confusió [163,164].

A més, l'Institut Català de la Salut, proveïdor de salut de 5.8 milions d'usuaris a Catalunya, està treballant en la validació de models d'IA i es troba en un desenvolupament incipient de diferents aplicacions [165,166]. No obstant això, actualment encara no n'hi ha cap d'integrat en la història clínica del pacient. Així doncs, i donat que en aquests estudis hi participen

diversos professionals sanitaris de la Catalunya Central, es pot haver generat un cert biaix i certs dubtes sobre el seu ús diari dins de les consultes.

A banda, l'ús de la intel·ligència activa també pot haver sigut un causant de la confusió respecte el seu ús i coneixement. Els professionals sanitaris d'atenció primària de l'ICS treballen amb l'estació clínica d'atenció primària (ECAP), on hi ha la història clínica del pacient informatitzada. Així doncs, l'ECAP és una eina de gestió clínica i administrativa que ofereix una visió integral de la persona al llarg del seu cicle vital i en els seus diferents processos de salut i malaltia en l'àmbit de l'atenció primària [167].

Dins de l'ECAP hi ha integrada la intel·ligència activa, un gestor clínic de dades que fusiona tota la informació del pacient i els seus problemes de salut, i suggereix per a cada pacient i problema de salut les millors opcions assistencial basades en l'evidència científica, presentades en un ordre lògic i temporal que s'adequa a la bona pràctica mèdica i d'infermeria [168]. Curiosament les inicials de la intel·ligència activa coincideixen amb les inicials de la IA però la intel·ligència activa no utilitza IA, fet que pot haver conduït al dubte o l'error, produint un cert biaix en els resultats de l'ús de la IA degut a la particularitat de l'àmbit d'estudi.

De fet, en relació amb la incertesa del coneixement i ús de la IA, dos de cada tres professionals ha indicat no haver rebut cap tipus de formació sobre la IA i nou de cada deu ha indicat que els agradaria rebre'n per estar més preparats. A més, els professionals han indicat no sentir-se preparats per la imminent implementació de la IA al seu dia a dia laboral considerant poca formació per poder utilitzar la IA amb seguretat a dins de la consulta. Aquests resultats concorden amb altres estudis on també es posa especial èmfasi en la necessitat i desig de formació que indiquen els professionals [43,48].

Per tant, cal obrir un debat per decidir quines estratègies son necessàries per poder revertir aquesta situació. En efecte, ja s'ha demostrat que entendre el perquè, el com i el on de la implementació de la IA així com obtenir la formació i educació necessària al respecte redueix la reticència i millora la seva implementació [169–171]. Tot i això, encara no esta clara quina formació necessària han de tenir els professionals sanitaris per poder utilitzar la IA a les

consultes, ni tampoc en quin moment acadèmic s'ha d'oferir, dubtant si s'hauria d'incloure directament en els plans docents dels graus acadèmics dels professionals o bé s'hauria d'oferir un cop ja estiguin al món laboral [158,172–178].

A més, cal també pactar en quin moment i quina informació es dona als professionals i estudiants per evitar discriminacions i assegurar un coneixement generalitzat profund i expert en l'àmbit sanitari. Per exemple, es podria actualitzar els plans d'estudis dels graus de medicina i infermeria directament des del departament d'educació per introduir-hi la IA. Alguns estudis on s'ha realitzat enquestes a estudiants de medicina sobre el coneixement i impacte de la IA han observat que, actualment, no hi ha formació específica ni reglada sobre la IA en els graus tot i haver-hi un gran interès i una clara necessitat [48,179].

En conseqüència, és clar que cal unificar esforços entre sanitat, educació i especialistes en IA per poder crear cursos, seminaris i material teòric amb informació homogeneïtzada i validada. D'aquesta manera, es podria evitar el sentiment de desconeixement entre els professionals i futurs professionals que sovint els porta a buscar informació a fonts de dades poc validades per falta de documents oficials. Fins i tot cal treballar conjuntament amb els mitjans de comunicació, ja que en alguns casos els professionals han indicat que el coneixement que tenen de la IA l'han adquirit a través dels mitjans sense un criteri basat en l'experiència o en informació teòrica validada [157].

Malgrat això, els professionals d'infermeria i medicina d'atenció primària consideren positiu l'impacte que pot tenir la IA al seu dia a dia dins de la consulta tot i que aquests nivells de positivitat són propers a la neutralitat. Aquests resultats concorden amb altres estudis on també s'ha observat un impacte positiu per part dels professionals sanitaris d'atenció primària degut a la facilitat d'ús, la disponibilitat en temps i ubicació, l'eficiència i qualitat del diagnòstic, l'augment de la velocitat i precisió dels processos, la reducció d'estigmes i prejudicis, la facilitat en la comunicació de l'equip assistencial, la capacitat d'assumir tasques repetitives, la possibilitat d'oferir una segona opinió als diagnòstics o avaluacions clíniques i la capacitat de fer un bon triatge dels pacients [46–48,81,137,155,157,180–182].

D'aquesta manera, s'ha mostrat esperança i predisposició de la IA en la possibilitat d'ajudar als professionals en la presa de decisions i, en conseqüència, en la millora de la salut de la població. Fins i tot, altres estudis han observat que tot i que especialistes d'atenció primària no veuen un impacte directe de la IA sobre les seves funcions, si que consideren que el suport d'aquestes aplicacions els pot ajudar a reduir el nombre de derivacions a especialistes [181]. Així com també han mostrat aquest sentiment positiu els estudiants de medicina, els qui creuen que podrà ser molt beneficiosa en un futur, però presenten més dubtes que els professionals a una possible substitució de la IA [47,48].

Coincidint amb els resultats de la població general, també s'observa que els professionals que tenen més coneixement i més predisposició a rebre coneixement sobre la IA perceben un impacte més positiu. En aquesta línia, altres estudis també han observat que fins i tot els professionals que ja utilitzen la IA consideren que l'impacte pot ser més positiu. Així doncs, és important que els professionals puguin tenir contacte amb aplicacions d'IA per poder conèixer com funciona i sobretot com els pot ajudar per poder reduir-ne la reticència [43,173,183]. A més, aquests estudis han posat èmfasi en la importància d'incloure els professionals dins el procés de creació i validació d'eines d'IA per oferir transparència, coneixement, ajustar-se a les necessitats de la pràctica clínica real i fer-los participants des de l'inici de l'aplicació [155].

D'altra banda, segons diversos estudis, no hi ha una preocupació general de l'impacte de la IA en el cost associat a la seva feina ni de la possibilitat que la IA els acabi substituint, especialment en la població més jove. De fet, no consideren que sigui una futura amenaça del seu lloc de treball, sinó que ho entenen com un suport que els podrà fer treballar més ràpid, més eficients i amb més fiabilitat. Existeix una visió general on el professional considera que la IA podria realitzar tasques més senzilles i administratives alliberant els professionals per fer tasques més complicades, així com generar informes de proves de manera automàtica per donar suport a la interpretació final del professional, però no podrà substituir mai la interacció humana i el tracte personalitzat [81,184].

Tot i això, en la literatura encara hi ha certa discrepància en l'aspecte de la substitució i alguns estudis indiquen que els professionals els preocupa el fet de poder ser substituïts per la IA tot i que el sentiment no és majoritari. Tanmateix, hi ha certa preocupació general d'un impacte de la IA negatiu en l'experiència dels nous professionals. És a dir, a la llarga, l'ús abusiu de la IA pot acabar generant una incapacitat per treballar de manera autònoma [157].

A més a més, tot i les grans oportunitats que pot oferir la IA, els professionals han indicat no sentir-se correctament preparats per la imminent aplicació de la IA, sobretot aquells qui han indicat no utilitzar-la al seu dia a dia. Existeixen certes preocupacions dels professionals i reptes específics de cara a la seva implementació com ara la necessitat de formació per guanyar coneixement i confiança, el risc de canviar la relació actual entre professional-pacient, la pèrdua d'autonomia del professional, i la necessitat d'adaptació i preparació de les infraestructures actuals.

Diversos estudis ja han indicat una falta de confiança en el funcionament intern de la IA, en la protecció de les dades, en la precisió, la transparència, la seguretat dels pacients, la privacitat i en la possible automatització total [39,46,81,157,185,186]. A més, degut al poc temps per pacient del que disposen els professionals d'atenció primària, els professionals estan d'acord en que les eines de IA han de ser fàcils d'utilitzar i ràpides per poder ser realment útils dins la consulta [157]. Tanmateix, també s'ha mostrat una necessitat de decidir qui pren el pes de la responsabilitat de la decisió de l'aplicació de IA així com la necessitat de definir en quins casos és apropiat el seu ús i en quins no, com per exemple en malalties molt poc prevalents i relacionades amb la salut mental [81].

Tot i que s'espera que la IA acabi afectant totes les especialitats de la medicina, actualment la radiologia és una de les especialitats on hi està tenint més impacte [38,173]. En aquest sentit, a la Catalunya Central i a altres territoris del món on hi ha una important falta de radiòlegs aquest impacte a la radiologia pot ser molt beneficiós. De fet, els especialistes d'atenció primària indiquen certa prioritat en disposar d'aplicacions d'IA per poder fer una primera lectura de les imatges immediata així com una interpretació dels resultats finals i un suport clínic en la presa de decisions.

A més, els professionals d'atenció primària son el primer contacte amb l'usuari i son els que tenen la capacitat de dirigir, en cas necessari, a l'usuari a un altre especialitat per ser tractat. Per això, és important que puguin disposar de eines de suport específiques de diferents especialitats per guanyar precisió, sentir-se més apoderats i reduir derivacions innecessàries a especialistes hospitalaris. A més, els professionals d'atenció primària tenen un temps molt curt per destinar a cada usuari i, tot i que de moment el sentiment no és general degut principalment al desconeixement, disposar d'eines els podria ajudar a guanyar agilitat i temps per dedicar-se a tasques més importants [157,187].

6.3 Validació externa de l'algorisme d'IA ChestEye

L'algorisme *ChestEye* ha demostrat una gran precisió a la interpretació de les imatges en comparació amb el *gold standard*, amb un valor de l'*accuracy* global de 0.95 (IC 95% 0.92; 0.98) . No obstant això, aquesta elevada precisió ha sigut deguda, en part, a la gran capacitat de l'algorisme per identificar correctament les imatges sense alteracions radiològiques, que eren la majoria de les imatges de la mostra. És a dir, l'algorisme presenta una bona capacitat de triatge entre imatges amb alteracions i sense alteracions. En canvi, no ha presentat bona capacitat per detectar certes patologies en les imatges, amb un valor de la sensibilitat global de 0.48 (IC 95% 0.30; 0.66) .

Aquest valor elevat de l'*accuracy* degut a l'elevada especificitat de l'algorisme és positiu ja que, com s'ha descrit anteriorment, la majoria de professionals sanitaris consideren que una de les utilitats i ajuda que han de donar els algorismes d'IA és la funció de triatge. D'aquesta manera, els professionals poden dedicar més temps a casos més complexos i no perdre temps en casos sense alteracions. Altres estudis també han demostrat valors d'especificitat elevats en validacions externes [117,188].

En relació a la sensibilitat, si es compara els resultats obtinguts amb altres validacions publicades, crida l'atenció la baixa sensibilitat de l'algorisme, perquè en general les sensibilitats dels algorismes solen estar entre 0.8 i 1 [89,189–192]. La principal explicació a

aquest resultat més baix és la falta de validacions externes a la literatura ja que actualment, la majoria dels estudis publicats desenvolupen validacions amb imatges tractades o seleccionades per conveniència i per tant, aquestes no son externes o independents a la mostra d'entrenament i això provoca una sobre estimació de les capacitats de l'algorisme [193].

De fet, tal i com es comenta a la introducció, un estudi publicat al 2019 a la revista *The Lancet* indica que només un 6% dels articles on es publica una validació d'algorismes d'IA en radiologia inclouen una correcta validació externa, i un estudi del 2021 indica que dels 100 productes d'IA en l'àmbit de la radiologia amb marca CE, només 36 presenten evidències de la seva eficàcia i només 18 presenten una validació de les seves capacitats i impacte en la salut del pacient (REF).

Actualment, existeix una problemàtica respecte les validacions externes ja que encara manca una definició objectiva i clara. Aquest fet fa que hi hagi dubtes en la diferència entre validació interna i validació externa, totes dues igual de necessàries però indicades en diferents moments del desenvolupament d'un algorisme d'IA. Així doncs, hi ha discrepàncies en l'ús de la paraula validació ja que alguns investigadors l'utilitzen per provar l'ajust durant l'entrenament i d'altres per validar l'algorisme final en comparació a un *gold standard* extern.

Aquest fet passa, principalment, perquè fins ara mancava la falta d'un clar marc regulador a nivell europeu per garantir el compliment legal de la IA, fixar requisits a complir en el desenvolupament, entrenament i validació. Això implicava poca transparència, confiança i acceptació dels algorismes d'IA. Amb la nova llei d'IA de la UE aprovada al 2024 caldrà més transparència de les aplicacions d'IA i més validacions i això obligarà a realitzar validacions finals amb dades prospectives o indicar les bases de protecció de dades (REF).

Ara bé, seguint amb l'anàlisi de la sensibilitat de l'algorisme, les diferències en la capacitat de l'algorisme desapareixen si es compara els resultats del present estudi exclusivament amb estudis on es desenvolupa correctament una validació externa. En aquests casos, la sensibilitat de l'algorisme sol prendre valors entre 0.5 i 1, inferiors a les validacions internes o resultats d'entrenament. Alguns estudis han presentat resultats de validació externa molt

prometedors, però les metodologies entre ells encara son molt diferents i això els fa poc generalitzables. A més, els resultats de les validacions externes encara depenen molt de l'entrenament del model i de la base de dades externa amb la que s'ha fet la validació [72,115,117,133,194–196].

Per exemple, hi ha validacions externes on la mostra, tot i ser externa al conjunt d'entrenament, no és seleccionada prospectivament del món real i per tant, les patologies trobades i la freqüència d'aquestes es veu alterada. A més, hi ha validacions externes on les imatges utilitzades només presenten una única patologia a la vegada, sent poc representatives de la pràctica clínica real, ja sigui perquè l'algorisme no detecta simultàniament diverses patologies d'un conjunt de les que ha estat entrenat o perquè només està entrenat per una única patologia [71,90,192,197,198].

En el cas del present estudi, amb imatges prospectives i recollides en l'àmbit real d'atenció primària, cal destacar els bons resultats de sensibilitat en patologies com els cables externals, la cardiomegàlia, les anomalies en costelles, els implants medul·lars, les vàlvules cardíacques o les marques intersticials. En canvi, no ha mostrat prou capacitat en detectar correctament patologies del mediastí, dels vasos o dels ossos, especialment en patologies prevalents com son l'emfisema pulmonar, l'atelectasi lineal o la prominència hilar, resultats que concorden amb altres estudis publicats on s'ha observat que els algorismes d'IA tenen més facilitat per detectar troballes més anòmales com la cardiomegàlia o els cables externals, i en canvi menys facilitat en anomalies més subtils com els nòduls o la consolidació [48,117,194,195] [72].

A més, també s'han observat problemes en la detecció de nòduls, coincidint amb troballes d'altres estudis [101,199,200]. Al tractar-se de patologies on la detecció a temps és molt important, els algorismes presenten molta sensibilitat en la detecció de nòduls generant falsos positius. Aquest fet pot ser, o no, beneficiós ja que s'ha demostrat que els falsos positius en diversos cribratges sovint poden ser perjudicials degut a un accés de proves, un augment a dosis de radiació innecessàries, alteracions a la salut mental dels pacients i més càrrega de treball per als sistemes sanitaris [200].

En el present estudi, la majoria dels casos l'algorisme ha confós nòduls amb arèoles del teixit mamari, i en altres estudis també s'ha reportat una confusió per grassa, pleura o malalties pulmonars intersticials [101,199,200]. Una possible explicació podria ser que l'entrenament de l'algorisme s'hagi desenvolupament majoritàriament amb imatges d'homes generant així un important biaix de gènere, així com amb una prevalença d'aquesta patologia molt superior a la real creant una sensibilitat excessiva [87,102].

En relació als biaixos dels algorismes d'IA, sovint parlem dels beneficis d'equitat i no d'aquests però no estan exempts de tenir certs biaixos inherents que cal corregir o, si més no, conèixer i controlar. El disseny i l'entrenament dels algorismes d'IA és un moment clau per evitar biaixos [201,202]. És a dir, si l'algorisme es desenvolupa i s'entrena amb dades sense eliminar biaixos, amb possibles factors de confusió, sense una bona representativitat de tota la població o amb una base de dades petita i homogènia, l'algorisme final segurament tindrà biaixos i fins i tot pot accentuar les desigualtats de la població com ja s'ha vist en diversos estudis [203–207].

És per això que alguns estudis ja han posat èmfasi en la importància de disposar d'una correcta base de dades per desenvolupar i entrenar el model, una base de dades representativa, heterogènia, que es pugui extrapolar a altre països, altres sexes, grups d'edat, origen ètnic, estats de salut, entre molts d'altres, per donar representativitat a totes les minories de la població. No obstant això, encara son molts estudis de validacions d'algorismes d'IA que no reporten les característiques sociodemogràfiques de la mostra i això limita la validació externa dels resultats [201,202,208].

Així doncs, cal seguir desenvolupant estudis per tal d'avaluar i millorar els algorismes d'IA i demostrar el seu gran potencial. D'aquesta manera, els professionals serien més conscients de l'impacte que pot representar el seu ús; es podria disposar sempre d'una segona opinió de la interpretació sense pràcticament temps d'espera, es podria millorar els fluxos de treball dels centres sanitaris, es podria millorar l'eficiència dels professionals reduint l'actual sobrecàrrega de feina i el temps dedicat a la interpretació, millorant la salut de la població [71,110].

És important, doncs, que els entrenaments dels algorismes es facin amb mostres representatives i ajustades a la realitat. Altres estudis on s'han desenvolupat validacions externes i s'han recollit també variables sociodemogràfiques com el sexe, l'edat, la ètnia i l'estat fumador han posat èmfasi en la importància d'entrenar amb mostres representatives per evitar diferències i discriminacions [209–212]. També, alguns estudis han indicat que els algorismes d'IA per la interpretació d'imatges de tòrax podrien ser més acurats si es pogués inserir informació prèvia de l'usuari, des de dades sociodemogràfiques fins a patologies prevalents o historial d'alteracions anteriors que podrien donar informació més completa per tal d'interpretar correctament les imatges [213].

A la pràctica clínica habitual sovint s'acompanya la interpretació de la radiografia amb la història clínica del pacient i en molts casos, cal veure imatges prèvies quan es tracta del seguiment d'una patologia concreta. De fet, en el present estudi s'ha corroborat que a molts informes del radiòleg el text de la interpretació de la radiografia indica que és un seguiment de la patologia del pacient i en aquests casos, la IA també necessitaria disposar de tota la informació per poder fer-ne la millor interpretació. Tal i com es demostra en diversos estudis, la interpretació d'una radiografia de tòrax millora significativament quan es conté informació de la història del pacient [213].

Un estudi recent desenvolupat a Chicago ha sigut dels primers en provar un algorisme d'IA per la lectura de tòrax capaç d'analitzar no només la radiografia actual, sinó també les anteriors. Així doncs, ofereix una interpretació més completa perquè té una contextualització i fins i tot pot oferir recomanacions per una millor recuperació. A més, a demostrat que la precisió és gairebé tan bona com la del radiòleg tot i la necessitat de seguir entrenant el model per millorar-ne les seves capacitats amb patologies no observades o amb molt poca freqüència [214].

En conjunt, i coincidint amb resultats d'altres articles, cal seguir realitzant validacions externes dels algorismes proposats d'IA per assegurar la seva capacitat en entorns clínics i reals i per tenir la oportunitat de poder identificar i rectificar possibles biaixos o errors dels algorismes. A

més, cal seguir treballant en la creació d'un marc regulador a nivell europeu per assegurar una validació adequada, acurada i homogènia de tots els algorismes d'IA.

7. LIMITACIONS

Els resultats de la present tesi doctoral presenten certes limitacions que cal esmentar per fer una correcta interpretació dels resultats. D'una banda, és important tenir en compte que les percepcions descrites de la present tesi son només consideracions hipotètiques ja que actualment la majoria dels professionals i usuaris no tenen una experiència d'ús real d'aplicacions d'IA en salut. És a dir, no només es tracta de percepcions i coneixements de l'ús de la IA sinó de l'actitud subjacent respecte aquestes eines [157].

Pel que fa a l'estudi dels usuaris, podria haver-hi un biaix d'edat ja que els conceptes de l'enquesta son més coneguts entre la població jove i a més, l'accés a través de QR no és utilitzat entre la població d'edats avançades. També podria haver-hi un biaix de selecció degut a que la majoria de la població de l'estudi eren persones que acudien a un centre de salut. Per últim, hi ha dues limitacions inherents de l'enquesta: l'idioma, perquè l'enquesta es va fer només en català, i la duplicitat, ja que en les enquestes a través de QR no es podia controlar si un usuari havia respost més d'una vegada.

Pel que fa a l'estudi dels professionals, podria haver-hi dos biaixos en la mostra degut a la particularitat del territori: hi ha un elevat percentatge de dones a la mostra, ja que dels professionals del territori d'estudi, el 89% dels professionals de medicina i el 65% dels professionals d'infermeria son dones, i hi ha una gran representació de zones rurals, generant poca validesa externa dels resultats.

A més, en relació als dos estudis anteriors, també hi ha una limitació de l'instrument de mesura utilitzat. En tots dos casos, es va utilitzar una enquesta validada a l'anglès, és a dir, no validada al català. A més, es va afegir preguntes dins de l'enquesta i, per tant, les dues enquestes utilitzades haurien de tornar a ser validades amb el canvi d'idioma i amb les noves preguntes afegides.

Per últim, pel que fa a la validació externa les principals limitacions son degudes a l'objectiu de no alterar la pràctica clínica habitual per poder fer una validació el màxim ajustada possible a la realitat. La primera d'aquesta limitacions és una mida de la mostra petita que no permet

validar bé totes les patologies de l'algorisme d'IA i que només s'ha obtingut d'un únic centre de salut. Al tractar-se de patologies en alguns casos molt poc prevalents, és difícil poder-les observar a les consultes reals en un període curt de temps. De fet, el gran volum de radiografies de l'estudi no tenen alteracions significatives. La segona limitació és que les interpretacions de les imatges no les feia sempre un únic professional, sinó que hi havia un grup de radiòlegs i a més, al voler provar l'algorisme en l'àmbit de l'atenció primària hi va haver una doble interpretació: una primera interpretació de la imatge per part del radiòleg i una segona interpretació de l'informe del radiòleg per part del metge de família i comunitària. La tercera limitació de l'estudi ha sigut no disposar d'informació sociodemogràfica com el sexe, l'edat o la ètnia dels pacients per poder avaluar la capacitat de l'algoritme en diferents grups poblacionals així com informació clínica anterior del pacient per poder contextualitzar bé cada imatge.

8. FUTURES LÍNIES D'INVESTIGACIÓ

A continuació s'indiquen futures línies d'investigació per poder entendre i aprofundir més sobre els resultats i conclusions de la present tesi doctoral:

- Desenvolupar estudis qualitius per poder conèixer la percepció i coneixements respecte l'ús de la IA com a eina de suport en salut de la població i dels professionals sanitaris. En el cas del present estudi s'han realitzat estudis transversals quantitius amb enquestes validades on les preguntes son en format tancat. Si es complementen aquests estudis amb estudis qualitius, es pot anar un pas més enllà.
- Recollir més mostra per ampliar la validació externa amb imatges de l'actual centre i d'altres centres i realitzar així una validació externa multicèntrica. Això permetria poder obtenir estimacions amb més precisió, poder observar patologies no observades en el present estudi i assegurar una millor representació de la població en els diferents centres.
- Recollir informació sociodemogràfica de la població que accedeix a participar en la validació de l'algorisme, així com informació clínica d'interès. D'aquesta manera es podria avaluar l'algorisme en cada un dels grups d'estudi per veure, per exemple, la capacitat en homes i la capacitat en dones.
- Si s'implementa l'algorisme dins del sistema, desenvolupar un estudi per avaluar l'impacte de l'algorisme dins el flux de treball dels centres sanitaris com els canvis en el nombre de derivacions des d'atenció primària a radiologia, els canvis en el temps d'espera per proves i per obtenció de resultats o els canvis en el nombre total de proves realitzades.

9. CONCLUSIONS

1. La majoria de la població indica que està familiaritzada amb el concepte d'IA tot i que aquest coneixement podria no ser profund sinó simples nocions bàsiques, remarcant la necessitat de posar èmfasi en la creació d'estratègies de formació per fer arribar el coneixement a tots els perfils poblacionals.
2. Existeix un sentiment envers l'aplicació de la IA en salut i especialment en radiologia lleugerament negatiu de la població. L'àmbit on la IA ha mostrat més preocupació ha sigut en la possible pèrdua de la interacció social, seguit de la necessitat d'estar informat i la desconfiança i la falta de responsabilitat de la IA. Per contra, l'àmbit on ha mostrat més esperança ha sigut en l'eficiència que pot oferir la IA.
3. D'una banda, i respecte la necessitat d'estar informat, la desconfiança i la responsabilitat de la IA, les persones amb nivells d'estudis inferiors i sense coneixements d'IA han mostrat puntuacions més negatives. D'altra banda, i respecte la desconfiança en l'aplicació de la IA i la seva falta de responsabilitat, les persones més grans de 65 anys i les persones de sexe femení són les que han indicat més preocupacions en aquests àmbits. Per tant, podrien ser els primers perfils on posar l'accent per revertir la situació de reticència i resistència al canvi.
4. La majoria dels professionals d'infermeria i medicina indiquen conèixer el concepte d'IA, mostren dubtes sobre si l'utilitzen o no a les consultes i indiquen una manca de formació en IA i un gran interès en rebre'n, accentuant així la importància i necessitat d'oferir informació validada i específica als professionals.
5. Els professionals han indicat una percepció positiva de l'impacte de la IA en el seu dia a dia laboral, especialment en la millora de les decisions clíniques i en la millora de la salut de la població. Aquesta percepció ha sigut major en els professionals amb coneixements i interès sobre la IA.
6. En canvi, existeix preocupació entre els professionals al no sentir-se correctament preparats per la imminent implementació de la IA en el dia a dia de les consultes, sobretot dels que indiquen que actualment no la utilitzen.

7. Respecte l'ús d'aplicacions d'IA en l'àmbit de la radiologia, els professionals d'infermeria i medicina d'atenció primària consideren prioritari disposar d'aquestes eines de suport per poder tenir interpretacions ràpides i resultats preliminars.
8. L'algorisme d'IA ChestEye per a la lectura de radiografies de tòrax ha demostrat ser útil i eficaç en la identificació d'imatges amb o sense alternacions radiològiques, i per tant podria ser una bona eina de cribratge. Tanmateix, cal una formació addicional per augmentar la capacitat diagnòstica d'algunes de les condicions analitzades així com l'entrenament per detectar noves patologies per les quals no ha estat entrenat com la broncopatia crònica, la bronquièctasi i l'engruiximent de la paret bronquial.
9. Les patologies amb major capacitat de detecció segons l'algorisme han sigut els cables externals, la cardiomegalia, les anomalies en les costelles, els implants medul·lars, les vàlvules cardíaques i les marques intersticials. En canvi, les patologies amb menor capacitat de detecció han sigut les patologies del mediastí, dels vasos o dels ossos com per exemple l'emfisema pulmonar, l'atelèctasi lineal o la prominència hilar.
10. Cal destacar la gran quantitat de falsos positius de l'algorisme en la detecció de nòduls, confonent aquests amb arèoles del teixit mamari.
11. Es podria modificar la nomenclatura d'alguna patologia identificada per l'algorisme per tal d'ajustar-se al vocabulari utilitzat dins el camp de la radiologia: les marques intersticials es podrien canviar per anomalia intersticial, la consolidació per condensació, l'esclerosi aòrtica per esclerosi valvular i la costella anormal per fractura costal.
12. Cal remarcar la importància de realitzar validacions externes en un entorn real, amb imatges reals, per tal d'obtenir resultats robusts ajustats a la realitat i no sobreestimats. Per això, cal disposar d'un marc regulador europeu on s'indiquin els requisits d'avaluació d'algorismes d'IA així com una definició acurada i validada de la correcta metodologia d'una validació externa.

10. BIBLIOGRAFIA

1. McCorduck P, Minsky M, Selfridge OG, Simon HA. History of artificial intelligence. IJCAI. 1977;951–4.
2. Haenlein M, Kaplan A. A Brief History of Artificial Intelligence: On the Past, Present, and Future of Artificial Intelligence. Calif Manage Rev. 2019;61(4):5–14.
3. Mondal B. Artificial intelligence: state of the art. Balas V, Kumar R, Srivastava R, editors. Vol. 172, Intelligence Systems Reference Library. Springer; 2020. 389-425p .
4. Berkeley. Artificial Intelligence: A Modern Approach, 4th US ed. [Internet]. 2023. Available from: <https://aima.cs.berkeley.edu/>
5. Bartneck C, Lütge C, Wagner A, Welsh S. What Is AI? SpringerBriefs in Ethics. An introduction to ethics in robotics and AI. 2021. 5-16 p.
6. REAL ACADEMIA ESPAÑOLA: Diccionario de la lengua española [Internet]. 23.^a ed. [versión 23.6 en línea]. 2023 [cited 2023 Nov 8]. Available from: <https://dle.rae.es>
7. Institut d’Estudis Catalans - Diec2 [Internet]. 2024 [cited 2024 Apr 8]. Available from: <https://dlc.iec.cat/>
8. OpenAI. ChatGPT [Internet]. ["Definition of artificial intelligence"]. 2023 [cited 2023 Nov 8]. Available from: <https://chat.openai.com>
9. Zhang C, Lu Y. Study on artificial intelligence: The state of the art and future prospects. J Ind Inf Integr. 2021;23.
10. García-Peñalvo FJ, Vázquez-Ingelmo A. What Do We Mean by GenAI? A Systematic Mapping of The Evolution, Trends, and Techniques Involved in Generative AI. Int J Interact Multimed Artif Intell. 2023;8(4):7–16.
11. stability.ai. Stable Diffusio 3 [Internet]. [cited 2024 Mar 25]. Available from: <https://stability.ai/news/stable-diffusion-3-research-paper>
12. Price WN, Cohen IG. Privacy in the age of medical big data. Vol. 25, Nature Medicine. 2019.
13. Albiol-Perarnau M, Belmonte IA. Blockchain en salud: transformando la seguridad y la gestión de datos clínicos. Atención Primaria. 2024;56(5).
14. Najafabadi MM, Villanustre F, Khoshgoftaar TM, Seliya N, Wald R, Muharemagic E. Deep learning applications and challenges in big data analytics. J Big Data. 2015;2(1).
15. Neyshabur B, Bhojanapalli S, McAllester D, Srebro N. Exploring generalization in deep learning. In: Advances in Neural Information Processing Systems. 2017.
16. Widrow B, M. A. Lehr. 30 years of adaptive neural networks: perceptron, Madaline,

- and backpropagation. *Proc IEEE*. 78(9):1415–42.
17. Goodfellow I, Y B, A C. Deep Learning. Vol. 29, *Nature*. 2019. 1-73 p.
 18. Zhao X, Qi S, Zhang B, Ma H, Qian W, Yao Y, et al. Deep CNN models for pulmonary nodule classification: Model modification, model integration, and transfer learning. *J Xray Sci Technol*. 2019;27(4).
 19. Roy D, Panda P, Roy K. Tree-CNN: A hierarchical Deep Convolutional Neural Network for incremental learning. *Neural Networks*. 2020;121.
 20. Au-Yong-Oliveira M, Lopes C, Soares F, Pinheiro G, Guimaraes P. What can we expect from the future the impact of Artificial Intelligence on Society. In: *Iberian Conference on Information Systems and Technologies, CISTI*. 2020.
 21. Raymond Geis J, Brady AP, Wu CC, Spencer J, Ranschaert E, Jaremko JL, et al. Ethics of artificial intelligence in radiology: Summary of the joint European and North American multisociety statement. *Radiology*. 2019;293(2).
 22. Bhandari M, Zeffiro T, Reddiboina M. Artificial intelligence and robotic surgery: Current perspective and future directions. Vol. 30, *Current Opinion in Urology*. 2020.
 23. Richardson JP, Smith C, Curtis S, Watson S, Zhu X, Barry B, et al. Patient apprehensions about the use of artificial intelligence in healthcare. *npj Digit Med*. 2021;4(1).
 24. Calderón CA, Perosanz JJI, Ortega-Mohedano F, Domínguez BG de G, Esteban FJF, Ramos MM, et al. *Ciencia de Datos en España*. Universidad de Salamanca. Salamanca. 2021. 1–85 p.
 25. Sánchez-Holgado P, Calderón CA, Blanco-Herrero D. Knowledge and attitudes of Spanish citizens towards big data and artificial intelligence. *Icono14*. 2022;20(1):1–25.
 26. Holgado PS, Calderón CA, Herrero DB. Conocimiento y percepción de la ciudadanía española sobre el big data y la inteligencia artificial. *Icono14*. 2022;20(1), 15.
 27. Maslej N, Fattorini L, Brynjolfsson E, Etchemendy J, Ligett K, Lyons T, et al. *The AI index 2023 Annual Report*. 2023. AI Index Steering Committee, Insititute for Human-Centered AI, Standford University, Standford; 2023
 28. Gillespie N, Lockey S, Curtis C, Pool J, Akbari A. *Trust in Artificial Intelligence: A Global Study*. Univ Queensl KPMG Aust. 2023;
 29. Araujo T, Helberger N, Kruikemeier S, de Vreese CH. In AI we trust? Perceptions about automated decision-making by artificial intelligence. *AI Soc*. 2020;35(3).
 30. Saheb T. Ethically contentious aspects of artificial intelligence surveillance: a social science perspective. *AI Ethics*. 2023;3(2).
 31. Madiega T. *EU Legislation in Progress: Artificial Inteligence act*. European Parliament;

- 2024.
32. Kaul V, Enslin S, Gross SA. History of artificial intelligence in medicine. Vol. 92, *Gastrointestinal Endoscopy*. 2020.
 33. Guo Y, Hao Z, Zhao S, Gong J, Yang F. Artificial intelligence in health care: Bibliometric analysis. *J Med Internet Res*. 2020;22(7).
 34. Solutions SHR cycle. Healthcare artificial intelligence outlook: Benefits, projected growth and challenges. [Internet]. The sis group. 2022. Available from: <https://thessigroup.com/blog/healthcare-artificial-intelligence-outlook-benefits-projected-growth-challenges/>
 35. GMI. Artificial intelligence in healthcare market size by application (medical imaging and diagnosis, drug discovery, therapy planning, hospital workflow, wearables, virtual assistants), COVID-19 impact analysis, regional outlook, application potential, price [Internet]. Global Market Insights. 2022. Available from: <https://www.gminsights.com/industry-analysis/healthcare-artificial-intelligence-market>
 36. Emani S, Rui A, Rocha HAL, Rizvi RF, Juaçaba SF, Jackson GP, et al. Physicians' Perceptions of and Satisfaction With Artificial Intelligence in Cancer Treatment: A Clinical Decision Support System Experience and Implications for Low-Middle–Income Countries. *JMIR Cancer*. 2022;8(2).
 37. Wu C, Xu H, Bai D, Chen X, Gao J, Jiang X. Public perceptions on the application of artificial intelligence in healthcare: a qualitative meta-synthesis. *BMJ Open*. 2023;13(1).
 38. Laï MC, Brian M, Mamzer MF. Perceptions of artificial intelligence in healthcare: Findings from a qualitative survey study among actors in France. *J Transl Med*. 2020 Jan 9;18(1):1–13.
 39. Jocelyn Chew HS, Achananuparp P. Perceptions and Needs of Artificial Intelligence in Health Care to Increase Adoption: Scoping Review. Vol. 24, *Journal of Medical Internet Research*. 2022.
 40. Fritsch SJ, Blankenheim A, Wahl A, Hetfeld P, Maassen O, Deffge S, et al. Attitudes and perception of artificial intelligence in healthcare: Across-sectional survey among patients. *Digit Heal*. 2022;8.
 41. Beets B, Newman TP, Howell EL, Bao L, Yang S. Surveying Public Perceptions of Artificial Intelligence in Health Care in the United States: Systematic Review. Vol. 25, *Journal of Medical Internet Research*. 2023.
 42. Sonawane A, Shah S, Pote S, He M. The application of artificial intelligence: perceptions from healthcare professionals. *Health Technol (Berl)*. 2023;
 43. Shinnars L, Aggar C, Stephens A, Grace S. Healthcare professionals' experiences and

- perceptions of artificial intelligence in regional and rural health districts in Australia. *Aust J Rural Health*. 2023;(June 2022):1–11.
44. Lai MC, Brian M, Mamzer MF. Perceptions of artificial intelligence in healthcare: Findings from a qualitative survey study among actors in France. *J Transl Med*. 2020 Jan 9;18(1).
 45. Abdullah R, Fakieh B. Health Care Employees' Perceptions of the Use of Artificial Intelligence Applications: Survey Study. *J Med Internet Res*. 2020 May 1;22(5).
 46. Castagno S, Khalifa M. Perceptions of Artificial Intelligence Among Healthcare Staff: A Qualitative Survey Study. *Front Artif Intell*. 2020 Oct 21;3.
 47. Swed S, Alibrahim H, Elkalagi NKH, Nasif MN, Rais MA, Nashwan AJ, et al. Knowledge, attitude, and practice of artificial intelligence among doctors and medical students in Syria: A cross-sectional online survey. *Front Artif Intell*. 2022;5.
 48. Kansal R, Bawa A, Bansal A, Trehan S, Goyal K, Goyal N, et al. Differences in Knowledge and Perspectives on the Usage of Artificial Intelligence Among Doctors and Medical Students of a Developing Country: A Cross-Sectional Study. *Cureus*. 2022;
 49. Gálvez M. Inteligencia artificial en radiología: ¿Seremos reemplazados por las máquinas? Vol. 23, *Revista Chilena de Radiología*. 2017.
 50. Mollura DJ, Azene EM, Starikovskiy A, Thelwell A, Iosifescu S, Kimble C, et al. White paper report of the rad-aid conference on international radiology for developing countries: Identifying challenges, opportunities, and strategies for imaging services in the developing world. Vol. 7, *Journal of the American College of Radiology*. 2010.
 51. Rimmer A. Radiologist shortage leaves patient care at risk, warns royal college. Vol. 359, *BMJ (Clinical research ed.)*. 2017.
 52. Royal College of Radiologists. Clinical radiology UK workforce census 2016 report. 2017.
 53. Royal College of Radiologists. Clinical radiology UK workforce census 2019 report. 2020.
 54. Zheng C, Deng X, Fu Q, Zhou Q, Feng J, Ma H, et al. Deep Learning-based Detection for COVID-19 from Chest CT using Weak Label. *IEEE Trans Med Imaging*. 2020;
 55. Wang X, Deng X, Fu Q, Zhou Q, Feng J, Ma H, et al. A Weakly-Supervised Framework for COVID-19 Classification and Lesion Localization from Chest CT. *IEEE Trans Med Imaging*. 2020;39(8).
 56. Elshennawy NM, Ibrahim DM. Deep-Pneumonia Framework Using Deep Learning Models Based on Chest X-Ray Images. *Diagnostics*. 2020;10(9).
 57. Degnan AJ, Ghobadi EH, Hardy P, Krupinski E, Scali EP, Stratchko L, et al. Perceptual

- and Interpretive Error in Diagnostic Radiology—Causes and Potential Solutions. *Acad Radiol.* 2019;26(6).
58. Graber ML, Franklin N, Gordon R. Diagnostic error in internal medicine. *Arch Intern Med.* 2005;165(13).
 59. Onder O, Yarasir Y, Azizova A, Durhan G, Onur MR, Ariyurek OM. Errors, discrepancies and underlying bias in radiology with case examples: a pictorial review. Vol. 12, *Insights into Imaging.* 2021.
 60. Kundel HL, Nodine CF, Carmody D. Visual scanning, pattern recognition and decision-making in pulmonary nodule detection. *Invest Radiol.* 1978;13(3).
 61. Patlas MN, Katz DS, Scaglione M. *Errors in Emergency and Trauma Radiology.* Springer Cham. 2007. 1;245
 62. Ruutiainen AT, Durand DJ, Scanlon MH, Itri JN. Increased Error Rates in Preliminary Reports Issued by Radiology Residents Working More Than 10 Consecutive Hours Overnight. *Acad Radiol.* 2013;20(3).
 63. Hanna TN, Shekhani H, Lamoureux C, Mar H, Nicola R, Sliker C, et al. Emergency Radiology Practice Patterns: Shifts, Schedules, and Job Satisfaction. *J Am Coll Radiol.* 2017;14(3).
 64. Pinto A, Reginelli A, Pinto F, Lo Re G, Midiri F, Muzj C, et al. Errors in imaging patients in the emergency setting. Vol. 89, *British Journal of Radiology.* 2016.
 65. Brady A, Laoide RÓ, McCarthy P, McDermott R. Discrepancy and error in radiology: Concepts, causes and consequences. Vol. 81, *Ulster Medical Journal.* 2012.
 66. Kim YW, Mansfield LT. Fool me twice: Delayed diagnoses in radiology with emphasis on perpetuated errors. *Am J Roentgenol.* 2014;202(3).
 67. Lodwick GS. Computer-aided diagnosis in radiology: A research plan. *Invest Radiol.* 1966;1(1).
 68. Novelline Ra and Squire LF. *Squire's fundamentals of radiology.* La Editorial U, editor. 2004.
 69. W H. *Learning radiology: Recognizing the basics.* Sciences EH, editor. 2015.
 70. Pezzotti W. *Interpretación de la radiografía de tórax: algo más que blanco y negro.* Nurs (Ed española). 2014;31(5).
 71. Hwang EJ, Park S, Jin KN, Kim JI, Choi SY, Lee JH, et al. Development and Validation of a Deep Learning–Based Automated Detection Algorithm for Major Thoracic Diseases on Chest Radiographs. *JAMA Netw Open.* 2019 Mar 22;2(3):e191095.
 72. Wu JT, Wong KCL, Gur Y, Ansari N, Karargyris A, Sharma A, et al. Comparison of Chest Radiograph Interpretations by Artificial Intelligence Algorithm vs Radiology Residents.

- JAMA Netw Open. 2020 Oct 1;3(10).
73. Santos Á, Solís P. Posición SERAM sobre la necesidad de informar la radiología simple. *Soc Española Radiol Médica*. 2015;(1):1–47.
 74. UNSCEAR. Sources and Effects of Ionizing Radiation Volume I: source. Vol. I, United Nations Scientific Committee on the Effects of Atomic Radiation. 2000. 1–654 p.
 75. Kasper D, Fauci A, Hauser S, Longo D, Jameson J, Loscalzo J. Harrison's principles of internal medicine. 19e ed. Vol. Vol. 1, No. 2. New York, NY, USA:: McGraw-hill; 2015.
 76. Dähnert W. Radiology Review Manual. *Nucl Med Commun*. 2011;32(10).
 77. Chassagnon G, Vakalopoulou M, Paragios N, Revel MP. Artificial intelligence applications for thoracic imaging. *Eur J Radiol*. 2020;123.
 78. Kim Y, Park JY, Hwang EJ, Lee SM, Park CM. Applications of artificial intelligence in the thorax: a narrative review focusing on thoracic radiology. Vol. 13, *Journal of Thoracic Disease*. 2021.
 79. Laino ME, Ammirabile A, Posa A, Cancian P, Shalaby S, Savevski V, et al. The applications of artificial intelligence in chest imaging of COVID-19 patients: A literature review. Vol. 11, *Diagnostics*. 2021.
 80. Mazurowski MA. Artificial Intelligence May Cause a Significant Disruption to the Radiology Workforce. *J Am Coll Radiol*. 2019;16(8).
 81. Goetz CM, Arnetz JE, Sudan S, Arnetz BB. Perceptions of virtual primary care physicians: A focus group study of medical and data science graduate students. *PLoS One*. 2020;15(12 December).
 82. Sociedad Española de Medicina de Familia y Comunitaria (semFYC). Programa de la especialidad de medicina familiar y comunitaria [Internet]. Ministerio de Sanidad y Consumo. 2005 [cited 2023 Sep 28]. Available from: <https://www.semfyec.es/wp-content/uploads/2016/09/Programa-Especialidad.pdf>
 83. Al Shammari M, Hassan A, AlShamlan N, Alotaibi S, Bamashmoos M, Hakami A, et al. Family medicine residents' skill levels in emergency chest X-ray interpretation. *BMC Fam Pract*. 2021;22(1).
 84. Schalekamp S, Klein WM, van Leeuwen KG. Current and emerging artificial intelligence applications in chest imaging: a pediatric perspective. Vol. 52, *Pediatric Radiology*. 2022.
 85. AI for radiology: an implementation guide. Diagnostic Image Analysis Group. <https://grand-challenge.org/aiforradiology/>. 2021.
 86. van Leeuwen KG, Schalekamp S, Rutten MJCM, van Ginneken B, de Rooij M. Artificial intelligence in radiology: 100 commercially available products and their scientific evidence. *Eur Radiol*. 2021;31(6).

87. Annarumma M, Withey SJ, Bakewell RJ, Pesce E, Goh V, Montana G. Automated triaging of adult chest radiographs with deep artificial neural networks. *Radiology*. 2019;291(1).
88. Yates EJ, Yates LC, Harvey H. Machine learning “red dot”: open-source, cloud, deep convolutional neural networks in chest radiograph binary normality classification. *Clin Radiol*. 2018;73(9).
89. Nam JG, Park S, Hwang EJ, Lee JH, Jin KN, Lim KY, et al. Development and Validation of Deep Learning-based Automatic Detection Algorithm for Malignant Pulmonary Nodules on Chest Radiographs. *Radiology*. 2019 Jan 1 [cited 2023 Sep 21];290(1):218–28.
90. Lakhani P, Sundaram B. Deep learning at chest radiography: Automated classification of pulmonary tuberculosis by using convolutional neural networks. *Radiology*. 2017;284(2).
91. Pasa F, Golkov V, Pfeiffer F, Cremers D, Pfeiffer D. Efficient Deep Network Architectures for Fast Chest X-Ray Tuberculosis Screening and Visualization. *Sci Rep*. 2019;9(1).
92. Qin ZZ, Sander MS, Rai B, Titahong CN, Sudrungrot S, Laah SN, et al. Using artificial intelligence to read chest radiographs for tuberculosis detection: A multi-site evaluation of the diagnostic accuracy of three deep learning systems. *Sci Rep*. 2019;9(1).
93. Rajpurkar P, O’Connell C, Schechter A, Asnani N, Li J, Kiani A, et al. CheXaid: deep learning assistance for physician diagnosis of tuberculosis using chest x-rays in patients with HIV. *NPJ Digit Med*. 2020;3(1):115.
94. Ardila D, Kiraly AP, Bharadwaj S, Choi B, Reicher JJ, Peng L, et al. End-to-end lung cancer screening with three-dimensional deep learning on low-dose chest computed tomography. *Nat Med*. 2019;25(6).
95. Baldwin DR, Gustafson J, Pickup L, Arteta C, Novotny P, Declerck J, et al. External validation of a convolutional neural network artificial intelligence tool to predict malignancy in pulmonary nodules. *Thorax*. 2020;75(4).
96. Shen S, Han SX, Aberle DR, Bui AA, Hsu W. An interpretable deep hierarchical semantic convolutional neural network for lung nodule malignancy classification. *Expert Syst Appl*. 2019;128.
97. Delzell DAP, Magnuson S, Peter T, Smith M, Smith BJ. Machine Learning and Feature Selection Methods for Disease Classification With Application to Lung Cancer Screening Image Data. *Front Oncol*. 2019;9.
98. Tu SJ, Wang CW, Pan KT, Wu YC, Wu C Te. Localized thin-section CT with radiomics feature extraction and machine learning to classify early-detected pulmonary nodules

- from lung cancer screening. *Phys Med Biol.* 2018;63(6).
99. Pehrson LM, Nielsen MB, Lauridsen CA. Automatic pulmonary nodule detection applying deep learning or machine learning algorithms to the LIDC-IDRI database: A systematic review. *Vol. 9, Diagnostics.* 2019.
 100. Li D, Vilmun BM, Carlsen JF, Albrecht-Beste E, Lauridsen CA, Nielsen MB, et al. The performance of deep learning algorithms on automatic pulmonary nodule detection and classification tested on different datasets that are not derived from LIDC-IDRI: A systematic review. *Vol. 9, Diagnostics.* 2019.
 101. Koo YH, Shin KE, Park JS, Lee JW, Byun S, Lee H. Extravalidation and reproducibility results of a commercial deep learning-based automatic detection algorithm for pulmonary nodules on chest radiographs at tertiary hospital. *J Med Imaging Radiat Oncol.* 2021;65(1).
 102. Sim Y, Chung MJ, Kotter E, Yune S, Kim M, Do S, et al. Deep convolutional neural network–based software improves radiologist detection of malignant lung nodules on chest radiographs. *Radiology.* 2020;294(1).
 103. Seah JCY, Tang JSN, Kitchen A, Gaillard F, Dixon AF. Chest radiographs in congestive heart failure: Visualizing neural network learning. *Radiology.* 2019;290(3).
 104. Taylor AG, Mielke C, Mongan J. Automated detection of moderate and large pneumothorax on frontal chest X-rays using deep convolutional neural networks: A retrospective study. *PLoS Med.* 2018;15(11).
 105. Walsh SLF, Calandriello L, Silva M, Sverzellati N. Deep learning for classifying fibrotic lung disease on high-resolution computed tomography: a case-cohort study. *Lancet Respir Med.* 2018;6(11).
 106. Christe A, Peters AA, Drakopoulos D, Heverhagen JT, Geiser T, Stathopoulou T, et al. Computer-Aided Diagnosis of Pulmonary Fibrosis Using Deep Learning and CT Images. *Invest Radiol.* 2019;54(10).
 107. Dorr F, Chaves H, Serra MM, Ramirez A, Costa ME, Seia J, et al. COVID-19 pneumonia accurately detected on chest radiographs with artificial intelligence. *Intell Med.* 2020;3–4.
 108. Kim JH, Kim JY, Kim GH, Kang D, Kim IJ, Seo J, et al. Clinical Validation of a Deep Learning Algorithm for Detection of Pneumonia on Chest Radiographs in Emergency Department Patients with Acute Febrile Respiratory Illness. *J Clin Med.* 2020 Jun 24;9(6):1981.
 109. Gang Nam J, Kim M, Park J, Jin Hwang E, Hyuk Lee J, Hee Hong J, et al. Development and validation of a deep learning algorithm detecting 10 common abnormalities on chest radiographs.
 110. Sung J, Park S, Lee SM, Bae W, Park B, Jung E, et al. Added value of deep learning-

- based detection system for multiple major findings on chest radiographs: A randomized crossover study. *Radiology*. 2021;299(2):450–9.
111. Li D, Pehrson LM, Lauridsen CA, Tøttrup L, Fraccaro M, Elliott D, et al. The added effect of artificial intelligence on physicians' performance in detecting thoracic pathologies on CT and chest X-ray: A systematic review. Vol. 11, *Diagnostics*. 2021.
 112. Kim DW, Jang HY, Kim KW, Shin Y, Park SH. Design characteristics of studies reporting the performance of artificial intelligence algorithms for diagnostic analysis of medical images: Results from recently published papers. *Korean J Radiol*. 2019;20(3).
 113. Liu X, Faes L, Kale AU, Wagner SK, Fu DJ, Bruynseels A, et al. A comparison of deep learning performance against health-care professionals in detecting diseases from medical imaging: a systematic review and meta-analysis. *Lancet Digit Heal*. 2019;1(6).
 114. Kriza C, Amenta V, Zenié A, Panidis D, Chassaigne H, Urbán P, et al. Artificial intelligence for imaging-based COVID-19 detection: Systematic review comparing added value of AI versus human readers. Vol. 145, *European Journal of Radiology*. 2021.
 115. Maiter A, Hocking K, Matthews S, Taylor J, Sharkey M, Metherall P, et al. Evaluating the performance of artificial intelligence software for lung nodule detection on chest radiographs in a retrospective real-world UK population. *BMJ Open*. 2023;13 (11).
 116. Kim C, Yang Z, Park SH, Hwang SH, Oh YW, Kang EY, et al. Multicentre external validation of a commercial artificial intelligence software to analyse chest radiographs in health screening environments with low disease prevalence. *Eur Radiol*. 2023 May 1];33(5):3501–9.
 117. van Beek EJR, Ahn JS, Kim MJ, Murchison JT. Validation study of machine-learning chest radiograph software in primary and emergency medicine. *Clin Radiol*. 2023;78(1).
 118. Oxipit. Study: AI found to reduce bias in Radiology Reports [Internet]. 2019 [cited 2023 Sep 22]. Available from: <https://oxipit.ai/news/study-ai-found-to-reduce-bias-in-radiology-reports/>
 119. Oxipit. Oxipit ChestEye secures medical device certification in Australia [Internet]. 2020 [cited 2023 Sep 22]. Available from: <https://oxipit.ai/news/oxipit-ai-medical-imaging-australia/>
 120. Comisión Europea. Comunicación de la comisión - Inteligencia artificial en Europa. COM(2018) 237 [Internet]. 2018. Available from: <https://eur-lex.europa.eu/legal-content/ES/TXT/?uri=COM%3A2018%3A237%3AFIN>
 121. Comisión Europea. Comunicación “Generar confianza en la Inteligencia Artificial centrada en el ser humano” (COM 168). Vol. COM (2019), *Diario Oficial de las Comunidades Europeas*. 2019.

122. Comisión Europea. Libro Blanco sobre la inteligencia artificial. 2020;1–32. Available from: https://ec.europa.eu/commission/sites/beta-political/files/political-guidelines-next-commission_es.pdf
123. Commission E. Regulation of the european parliament and of the council laying down harmonised rules on artificial intelligence and amending certain union legislative acts. 2024;1(1):98–101.
124. CIVICAi. Portada - CIVIC Ai [Internet]. [cited 2024 Apr 8]. Available from: <https://civicai.cat/>
125. Terry N. NP Regulatory disruption and arbitrage in health-care data protection. *Yale J Heal Pol’y L Ethics*. 2017;17:143–208.
126. Terry N. Existential challenges for healthcare data protection in the United States. *Ethics, Med Public Heal*. 2017;3(1).
127. General Data Protection Regulation. Commission Regulation 2016/679 of the European Parliament and of the Council of 27 April 2016 on the Protection of Natural Persons with regard to the Processing of Personal Data and on the Free Movement of such Data, and Repealing Directive, 95/46/EC. OJ [Internet]. 2016;(L 119):1, 34. Available from: <http://www.privacy-regulation.eu/en/article-4definitions-GDPR.htm>.
128. Ministerio de Sanidad G de E. ASPECTOS RELEVANTES Informe del SNS 2020-2021. BOE. 2021;
129. Myllyaho L, Raatikainen M, Männistö T, Mikkonen T, Nurminen JK. Systematic literature review of validation methods for AI systems. *J Syst Softw*. 2021;181.
130. Domalpally A, Channa R. Real-world validation of artificial intelligence algorithms for ophthalmic imaging. Vol. 3, *The Lancet Digital Health*. 2021.
131. Altman DG, Royston P. What do we mean by validating a prognostic model? In: *Statistics in Medicine*. 2000.
132. Nam JG, Kim M, Park J, Hwang EJ, Lee JH, Hong JH, et al. Development and validation of a deep learning algorithm detecting 10 common abnormalities on chest radiographs. *Eur Respir J*. 2021;57(5).
133. Sjoding MW, Taylor D, Motyka J, Lee E, Co I, Claar D, et al. Deep learning to detect acute respiratory distress syndrome on chest radiographs: a retrospective study with external validation. *Lancet Digit Heal*. 2021;3(6).
134. Lee JH, Sun HY, Park S, Kim H, Hwang EJ, Goo JM, et al. Performance of a deep learning algorithm compared with radiologic interpretation for lung cancer detection on chest radiographs in a health screening population. *Radiology*. 2020 Dec 1;297(3):687–96.
135. Ongena YP, Haan M, Yakar D, Kwee TC. Patients’ views on the implementation of

- artificial intelligence in radiology: development and validation of a standardized questionnaire. *Eur Radiol* 2019 30(2):1033–40.
136. De Regge M, Beirão G, Den Ambtman A, De Pourcq K, Dias J, Kandampully J. Health care technology adaption for elderly: does the family matter ? *Quis*. 2017;15.
 137. Shinnars L, Aggar C, Grace S, Smith S. Exploring healthcare professionals' perceptions of artificial intelligence: Validating a questionnaire using the e-Delphi method. *Digit Heal*. 2021 Jan 23;7:20552076211003430.
 138. Turja T, Rantanen T, Oksanen A. Robot use self-efficacy in healthcare work (RUSH): development and validation of a new measure. *AI Soc*. 2019 Mar 14 [cited 2024 Apr 8];34(1):137–43.
 139. Vargas E, Olivé M, Jurado JM, Boadas J, Riera I, Ureña M. Espai interconsulta-consulta virtual: millorant la integració entre nivells assistencials. *Ann Med*. 2014;97(2), 56–58.
 140. Wittal C, Hammer D, Klein F, Rittchen J. Perception and Knowledge of Artificial Intelligence in Healthcare, Therapy and Diagnostics: A Population-Representative Survey. *J Biotechnol Biomed*. 2023;06(02).
 141. Calderón C, Perosanz J, Ortega-Mohedano F, Domínguez B de G, Esteban F, Ramos M. Ciencia de Datos en España. Universidad de Salamanca. España. 2021. 1–3 p
 142. Zhang Z, Citardi D, Wang D, Genc Y, Shan J, Fan X. Patients' perceptions of using artificial intelligence (AI)-based technology to comprehend radiology imaging data. *Health Informatics J*. 2021;27(2).
 143. Ahmad Z, Rahim S, Zubair M, Abdul-Ghafar J. Artificial intelligence (AI) in medicine, current applications and future role with special emphasis on its potential and promise in pathology: present and future impact, obstacles including costs and acceptance among pathologists, practical and philosophical considerations. A comprehensive review. Vol. 16, *Diagnostic Pathology*. 2021.
 144. Haan M, Ongena YP, Hommes S, Kwee TC, Yakar D. A Qualitative Study to Understand Patient Perspective on the Use of Artificial Intelligence in Radiology. *J Am Coll Radiol*. 2019;16(10).
 145. Bahakeem BH, Alobaidi SF, Alzahrani AS, Alhasawi R, Alzahrani A, Alqahtani W, et al. The General Population's Perspectives on Implementation of Artificial Intelligence in Radiology in the Western Region of Saudi Arabia. *Cureus*. 2023;
 146. Ploug T, Sundby A, Moeslund TB, Holm S. Population Preferences for Performance and Explainability of Artificial Intelligence in Health Care: Choice-Based Conjoint Survey. *J Med Internet Res*. 2021;23(12).
 147. Lennartz S, Dratsch T, Zopfs D, Persigehl T, Maintz D, Hokamp NG, et al. Use and control of artificial intelligence in patients across the medical workflow: Single-center questionnaire study of patient perspectives. Vol. 23, *Journal of Medical Internet*

- Research. 2021.
148. Yakar D, Ongena YP, Kwee TC, Haan M. Do People Favor Artificial Intelligence Over Physicians? A Survey Among the General Population and Their View on Artificial Intelligence in Medicine. *Value Heal.* 2022;25(3).
 149. Miller S, Gilbert S, Virani V, Wicks P. Patients↔ utilization and perception of an artificial intelligence↓based symptom assessment and advice technology in a British primary care waiting room: Exploratory pilot study. *JMIR Hum Factors.* 2020;7(3).
 150. Comisió Europea. European Commission. Digital inclusion. <https://digitalstrategy.ec.europa.eu/en/policies/digital-inclusion> . 2021.
 151. Fontelo P, Liu F. A review of recent publication trends from top publishing countries. *Syst Rev.* 2018;7(1).
 152. Haleem S, Albar NH, Al fahad MS, AlWasem HO. Knowledge, Awareness, and Perception of COVID-19 and Artificial Intelligence: A Cross-Sectional Study Among the Population in Saudi Arabia. *Cureus.* 2023;
 153. Fast E, Horvitz E. Long-Term Trends in the Public Perception of Artificial Intelligence. 31st AAAI Conf Artif Intell AAAI 2017. 2016 Sep 16;963–9.
 154. Kaminski J. Diffusion of Innovation Theory Canadian Journal of Nursing Informatics. *Can J Nurs Informatics.* 2012;6(2).
 155. Liyanage H, Liaw ST, Jonnagaddala J, Schreiber R, Kuziemy C, Terry AL, et al. Artificial Intelligence in Primary Health Care: Perceptions, Issues, and Challenges: Primary Health Care Informatics Working Group Contribution to the Yearbook of Medical Informatics 2019. *Yearb Med Inform.* 2019;28(1).
 156. Scheetz J, Rothschild P, McGuinness M, Hadoux X, Soyer HP, Janda M, et al. A survey of clinicians on the use of artificial intelligence in ophthalmology, dermatology, radiology and radiation oncology. *Sci Rep.* 2021;11(1).
 157. Buck C, Doctor E, Hennrich J, Jöhnk J, Eymann T. General Practitioners' Attitudes Toward Artificial Intelligence–Enabled Systems: Interview Study. *J Med Internet Res.* 2022;24(1).
 158. Chen M, Zhang B, Cai Z, Seery S, Gonzalez MJ, Ali NM, et al. Acceptance of clinical artificial intelligence among physicians and medical students: A systematic review with cross-sectional survey. *Front Med.* 2022 Aug 31;9.
 159. Frey CB, Osborne MA. The future of employment: How susceptible are jobs to computerisation? *Technol Forecast Soc Change.* 2017;114.
 160. Smith A, Page D. Public Predictions for the Future of Workforce Automation: Full Report. Pew Research Center. 2016.
 161. Smith A, Anderson M. Automation in Everyday Life. *Pew Res Cent.* 2017;

162. Vidal-Alaball J, Seguí FL, Domingo JLG, Mateo GF, Valmaña GS, Ruiz-Comellas A, et al. Primary Care Professionals' Acceptance of Medical Record-Based, Store and Forward Provider-to-Provider Telemedicine in Catalonia: Results of a Web-Based Survey. *Int J Environ Res Public Health*. 2020 Jun 1;17(11):1–13.
163. Escalé-Besa A, Fuster-Casanovas A, Börve A, Yélamos O, Fustà-Novell X, Rafat ME, et al. Using Artificial Intelligence as a Diagnostic Decision Support Tool in Skin Disease: Protocol for an Observational Prospective Cohort Study. *JMIR Res Protoc*. 2022 Aug 1;11(8).
164. Vidal-Alaball J, Fibla DR, Zapata MA, Marin-Gomez FX, Fernandez OS. Artificial Intelligence for the Detection of Diabetic Retinopathy in Primary Care: Protocol for Algorithm Development. *JMIR Res Protoc*. 2019 Feb 1;8(2).
165. Coma E, Miró Q, Medina M, Marin-Gomez FX, Cos X, Benítez M, et al. Association between the reduction of face-to-face appointments and the control of patients with type 2 diabetes mellitus during the Covid-19 pandemic in Catalonia. *Diabetes Res Clin Pract*. 2021 Dec 1;182.
166. Programa per la promoció i desenvolupament de la Intel·ligència Artificial al Sistema de Salut de Catalunya IA.
167. Departament de Salut. Generalitat de Catalunya. Estació Clínica d'Atenció Primària (ECAP). <https://salutweb.gencat.cat/ca/ambits-actuacio/linies/tic/sistemes-informacio/gestio-assistencial/ecap/>. 2021.
168. Departament de Salut. Generalitat de Catalunya. L'ICS incorpora la intel·ligència activa a la pràctica assistencial. <https://www.icsgirona.cat/ca/noticies/primaria/193>. 2024.
169. Health Education England. The Topol Review: Preparing the healthcare workforce to deliver the digital future. An independent report on behalf of the Secretary of State for Health and Social Care. *Nhs*. 2019;(February):102.
170. Teasley R, Jordan M, Sangtani V. the Human Side of Technology Project Performance: Effects of Satisfaction, Perceived Technology Policy, Task Significance and Training. *Acad Inf Manag Sci J*. 2012;15(2).
171. Nash DM, Thorpe C, Brown JB, Kueper JK, Rayner J, Lizotte DJ, et al. Perceptions of Artificial Intelligence Use in Primary Care: A Qualitative Study with Providers and Staff of Ontario Community Health Centres. *J Am Board Fam Med*. 2023;36(2).
172. Civaner MM, Uncu Y, Bulut F, Chalil EG, Tatli A. Artificial intelligence in medical education: a cross-sectional needs assessment. *BMC Med Educ*. 2022 Dec 1;22(1):1–9.
173. Sit C, Srinivasan R, Amlani A, Muthuswamy K, Azam A, Monzon L, et al. Attitudes and perceptions of UK medical students towards artificial intelligence and radiology: a

- multicentre survey. *Insights Imaging*. 2020 Dec 1;11(1):1–6.
174. Wartman SA, Donald Combs C. Medical Education Must Move From the Information Age to the Age of Artificial Intelligence. *Acad Med*. 2018;93(8):1107–9.
 175. Liaw W, Kakadiaris IA. Artificial Intelligence and Family Medicine: Better Together. *Fam Med*. 2020;52(1):8–10.
 176. Katznelson G, Gerke S. The need for health AI ethics in medical school education. *Adv Health Sci Educ Theory Pract*. 2021 Oct 1;26(4):1447–58.
 177. Lee J, Wu AS, Li D, Kulasegaram KM. Artificial Intelligence in Undergraduate Medical Education: A Scoping Review. *Acad Med*. 2021 Nov 1;96(11S):S62–70.
 178. Wang Z, Paranjape R. A signal processing application for evaluating self-monitoring blood glucose strategies in a software agent model. *Comput Methods Programs Biomed*. 2015;120(2).
 179. McCoy LG, Nagaraj S, Morgado F, Harish V, Das S, Celi LA. What do medical students actually need to know about artificial intelligence? Vol. 3, *npj Digital Medicine*. 2020.
 180. Nadarzynski T, Miles O, Cowie A, Ridge D. Acceptability of artificial intelligence (AI)-led chatbot services in healthcare: A mixed-methods study. *Digit Heal*. 2019 Aug 1;5.
 181. Sangers TE, Wakkee M, Moolenburgh FJ, Nijsten T, Lugtenberg M. Towards successful implementation of artificial intelligence in skin cancer care: a qualitative study exploring the views of dermatologists and general practitioners. *Arch Dermatol Res*. 2023;315(5).
 182. Eastwood KW, May R, Andreou P, Abidi S, Abidi SSR, Loubani OM. Needs and expectations for artificial intelligence in emergency medicine according to Canadian physicians. *BMC Health Serv Res*. 2023;23(1).
 183. Jarva E, Mikkonen K, Andersson J, Tuomikoski AM, Kääriäinen M, Meriläinen M, et al. Aspects associated with health care professionals' digital health competence development – a qualitative study. *Finnish J eHealth eWelfare*. 2022;14(1).
 184. Doraiswamy PM, Blease C, Bodner K. Artificial intelligence and the future of psychiatry: Insights from a global physician survey. *Artif Intell Med*. 2020;102.
 185. Shinnars L, Aggar C, Stephens A, Grace S. Healthcare professionals' experiences and perceptions of artificial intelligence in regional and rural health districts in Australia. *Aust J Rural Health*. 2023 Dec 1;31(6):1203–13.
 186. Held LA, Wewetzer L, Steinhäuser J. Determinants of the implementation of an artificial intelligence-supported device for the screening of diabetic retinopathy in primary care – a qualitative study. *Health Informatics J*. 2022;28(3).
 187. Police RL, Foster T, Wong KS. Adoption and use of health information technology in physician practice organisations: Systematic review. Vol. 18, *Informatics in Primary*

- Care. 2011.
188. Blake SR, Das N, Tadepalli M, Reddy B, Singh A, Agrawal R, et al. Using Artificial Intelligence to Stratify Normal versus Abnormal Chest X-rays: External Validation of a Deep Learning Algorithm at East Kent Hospitals University NHS Foundation Trust. *Diagnostics* 2023, Vol 13, Page 3408. 2023 Nov 9;13(22):3408.
 189. Futoma J, Simons M, Panch T, Doshi-Velez F, Celi LA. The myth of generalisability in clinical research and machine learning in health care. *Lancet Digit Heal*. 2020 Sep 1 ;2(9):e489–92.
 190. Park SH. Diagnostic case-control versus diagnostic cohort studies for clinical validation of artificial intelligence algorithm performance. Vol. 290, *Radiology*. 2019.
 191. JT W, KCL W, Y G, N A, A K, A S, et al. Comparison of Chest Radiograph Interpretations by Artificial Intelligence Algorithm vs Radiology Residents. *JAMA Netw open*. 2020 Oct 1;3(10).
 192. Park S, Lee SM, Lee KH, Jung KH, Bae W, Choe J, et al. Deep learning-based detection system for multiclass lesions on chest radiographs: comparison with observer readings. *Eur Radiol*. 2020;30(3).
 193. Park SH, Han K. Methodologic guide for evaluating clinical performance and effect of artificial intelligence technology for medical diagnosis and prediction. Vol. 286, *Radiology*. 2018.
 194. Hwang EJ, Nam JG, Lim WH, Park SJ, Jeong YS, Kang JH, et al. Deep learning for chest radiograph diagnosis in the emergency department. *Radiology*. 2019;293(3):573–80.
 195. Rajpurkar P, Irvin J, Ball RL, Zhu K, Yang B, Mehta H, et al. Deep learning for chest radiograph diagnosis: A retrospective comparison of the CheXNeXt algorithm to practicing radiologists. Sheikh A, editor. *PLOS Med*. 2018 Nov 20 ;15(11):e1002686.
 196. Rajpurkar P, Irvin J, Zhu K, Yang B, Mehta H, Duan T, et al. CheXNet: Radiologist-Level Pneumonia Detection on Chest X-Rays with Deep Learning. 2017; 1711.05225
 197. Nam JG, Park S, Hwang EJ, Lee JH, Jin KN, Lim KY, et al. Development and validation of deep learning-based automatic detection algorithm for malignant pulmonary nodules on chest radiographs. *Radiology*. 2019;290(1).
 198. Hwang EJ, Park S, Jin KN, Kim JI, Choi SY, Lee JH, et al. Development and Validation of a Deep Learning-based Automatic Detection Algorithm for Active Pulmonary Tuberculosis on Chest Radiographs. *Clin Infect Dis*. 2019;69(5).
 199. Salas Pérez R, Teixidó Vives M, Picas Cutrina E, Romero Novo I. Diferentes aspectos de las calcificaciones mamarias. *Imagen Diagnóstica*. 2013 Jul 1;4(2):52–7.
 200. Maiter A, Hocking K, Matthews S, Taylor J, Sharkey M, Metherall P, et al. Original research: Evaluating the performance of artificial intelligence software for lung

- nodule detection on chest radiographs in a retrospective real-world UK population. *BMJ Open*. 2023 Nov 8;13(11):77348.
201. Cirillo D, Catuara-Solarz S, Morey C, Guney E, Subirats L, Mellino S, et al. Sex and gender differences and biases in artificial intelligence for biomedicine and healthcare. Vol. 3, *npj Digital Medicine*. 2020.
 202. Abràmoff MD, Tarver ME, Loyo-Berrios N, Trujillo S, Char D, Obermeyer Z, et al. Considerations for addressing bias in artificial intelligence for health equity. *npj Digit Med*. 2023;6(1).
 203. Leslie D, Mazumder A, Peppin A, Wolters MK, Hagerty A. Does “aI” stand for augmenting inequality in the era of covid-19 healthcare? *BMJ*. 2021;372.
 204. Obermeyer Z, Powers B, Vogeli C, Mullainathan S. Dissecting racial bias in an algorithm used to manage the health of populations. *Science* (80-). 2019;366(6464).
 205. Obermeyer Z, Nissan R, Stern M, Eaneff S, Bembeneck EJ, Mullainathan S. *Algorithmic Bias Playbook*. Chicago Booth - Cent Appl Artif Intell. 2021;
 206. Abràmoff MD, Tobey D, Char DS. Lessons Learned About Autonomous AI: Finding a Safe, Efficacious, and Ethical Path Through the Development Process. *Am J Ophthalmol*. 2020;214.
 207. Char DS, Abràmoff MD, Feudtner C. Identifying Ethical Considerations for Machine Learning Healthcare Applications. *Am J Bioeth*. 2020;20(11).
 208. Celi LA, Cellini J, Charpignon ML, Dee EC, Dernoncourt F, Eber R, et al. Sources of bias in artificial intelligence that perpetuate healthcare disparities—A global review. *PLOS Digit Heal*. 2022 Mar 31;1(3):e0000022.
 209. Gichoya JW, Banerjee I, Bhimireddy AR, Burns JL, Celi LA, Chen LC, et al. AI recognition of patient race in medical imaging: a modelling study. *Lancet Digit Heal*. 2022;4(6).
 210. Seyyed-Kalantari L, Zhang H, McDermott MBA, Chen IY, Ghassemi M. Underdiagnosis bias of artificial intelligence algorithms applied to chest radiographs in under-served patient populations. *Nat Med*. 2021;27(12).
 211. Wiens J, Creary M, Sjoding MW. AI models in health care are not colour blind and we should not be either. Vol. 4, *The Lancet Digital Health*. 2022.
 212. Char DS, Shah NH, Magnus D. Implementing Machine Learning in Health Care — Addressing Ethical Challenges. *N Engl J Med*. 2018;378(11).
 213. Potchen EJ, Gard JW, Lazar P, Lahaie P, Andary M. The effect of clinical history data on chest film interpretation: direction or distraction. *Invest Radiol*. 1979;14, 404.
 214. Huang J, Neill L, Wittbrodt M, Melnick D, Klug M, Thompson M, et al. Generative Artificial Intelligence for Chest Radiograph Interpretation in the Emergency Department. *JAMA Netw Open*. 2023 Oct 2;6(10):e2336100–e2336100.

11. ANNEXES

Annex I Enquesta original ciutadania

Artificial Intelligence in Medicine

Questionnaire administered to the LISS panel

Version 1.0

date	May 2021
author	Marika de Bruijne T: +31 13 206 3570 E: marika.de.debruijne@centerdata.nl P.O.Box 90153, 5000 LE Tilburg, the Netherlands Professor de Moorplein 524-525, Building I, 3 rd floor

© CentERdata, Tilburg, 2021

All rights reserved. No part of this publication may be reproduced, stored in a retrieval system, or transmitted, in any form or by any means, electronic, mechanical, photocopying, recording or otherwise, without the prior written permission of the publisher.



Table of contents

1	Project description.....	2
2	Introduction	3
3	Codebook.....	4



1 Project description

Title: Artificial Intelligence in Medicine

Data file: qv20a

Funding source(s): ODISSEI

Researcher(s): Marieke Haan (University of Groningen); Yfke Ongena (University of Groningen)

Project description: Artificial Intelligence (AI) is being incorporated into medicine to improve patient care. This research aims to map the general populations' view on the implementation of AI in medicine. Furthermore, different response scales (agree-disagree versus construct-specific scales) and their effect on data quality are experimentally tested. This research aims to answer questions that are relevant for survey practice, general medicine, and the society as a whole.

Sample: A random subsample of panel members aged 16 years or older

Overview of the response:

Selected number of household members:	3,117	(100.0%)
Non-response:	675	(21.7%)
Response:	2,442	(78.3%)
Complete:	2,411	(77.4%)
Incomplete:	31	(1.0%)

Date of data collection: April 2020

Titel: Kunstmatige Intelligentie in de Geneeskunde

Databestand: qv20a

Financiering: ODISSEI

Onderzoeker(s): Marieke Haan (Rijksuniversiteit Groningen); Yfke Ongena (Rijksuniversiteit Groningen)

Projectbeschrijving: Kunstmatige Intelligentie (KI) wordt steeds meer gebruikt in de geneeskunde om de patiëntenzorg te verbeteren. In dit onderzoek wordt de mening van de algemene bevolking over de implementatie van KI in de geneeskunde in kaart gebracht. Bovendien worden verschillende antwoordschalen (eens-oneens versus constructspecifieke schalen) en hun effecten op datakwaliteit experimenteel getoets. Dit onderzoek wil medische en methodologische vragen beantwoorden die relevant zijn voor de onderzoekspraktijk, de geneeskunde en de samenleving als geheel.

Steekproef: Een random subsample van panelleden die 16 jaar en ouder zijn

Responsoverzicht:

Selectie aantal leden huishouden:	3.117	(100,0%)
Nonrespons:	675	(21,7%)
Respons:	2.442	(78,3%)
Compleet:	2.411	(77,4%)
Incompleet:	31	(1,0%)

Datum van dataverzameling: april 2020



2 Introduction

In April 2020, the questionnaire “Kunstmatige intelligentie” (Artificial intelligence) was fielded in the LISS panel¹.

The questionnaire was presented to 3,117 panel members, and 2,411 respondents fully completed the questionnaire (response percentage 77%).

¹ Reference to LISS panel data in text:

In this paper we make use of data of the LISS (Longitudinal Internet studies for the Social Sciences) panel administered by CentERdata (Tilburg University, The Netherlands).



3 Codebook

This codebook contains the questionnaire as administered to the LISS panel.

- Variable names: The variable names are printed in **bold** and correspond to the names in the dataset.
- The questionnaire routing is printed in *italics* for each variable concerned.
 - = means "is equal to"
 - != means "is not equal to"
 - <= means "smaller or equal to"
 - >= means "larger or equal to"
- If there were buttons on a page for "I don't know, "I refuse to say" or "Not applicable", this will be indicated at the top at "Page" (when the button is clicked, the corresponding value is written to all variables that were present on the page)
- *text*: answer box (no limit to the length of the answer).
- *string*: answer box allowing a certain number of characters (standard is 255).
- *empty*: questions could be left unanswered
- Numerical variables: If ranges were used, these are printed in *italics* in the codebook if the respondent could not see them on the screen. If the respondent was able to see the ranges, the variables are printed in normal letter. *Integer*: If a question is not subject to any limit (integer), no range is indicated.
- 'Fills' (variable text) are given between straight brackets [].
- Variables in between curly brackets {} are not part of the dataset, but the corresponding questions or text were part of the questionnaire.

nomem_encr²

Number of the household member encrypted

qv20a_m

Year and month of the field work period

constructed variable

qv20a001

Mobile layout

0. No

1. Yes

² Use nomem_encr as key variable for merging with other LISS panel studies from the LISS Panel Data Archive: <https://www.dataarchive.lissdata.nl>



random variable

qv20a002

Which group was the respondent randomly assigned to?

1. group 1; questions with agree-disagree answer scale, fully labeled
2. group 2; questions with agree-disagree answer scale, end-point labeled
3. group 3; questions with construct-specific answer scale, fully labeled
4. group 4; questions with construct-specific answer scale, end-point labeled

Note: For all groups, the answer scales were displayed horizontally on the screen.

Page 1

{intro}

Developments in the field of computers go very fast. In the future, many medical procedures can be (partly) performed independently by a computer. This assistance from computers is also referred to as '**artificial intelligence**'.

We would like to know what you think about this use of (partly) independently operating computers when performing medical tasks.

First, a number of statements about artificial intelligence in the hospital follow.

Page 2

if qv20a002 = 1

qv20a003 - qv20a006

Please choose the answer that best reflects your opinion.

qv20a003 I find using computers to perform medical tasks a bad idea

qv20a004 I find using computers to perform medical tasks safe

qv20a005 I find using computers to perform medical tasks useful

qv20a006 I find using computers to perform medical tasks nifty

1. Strongly disagree
2. Disagree
3. Neither agree nor disagree
4. Agree
5. Strongly agree



Page 3

if qv20a002 = 1

qv20a007 - qv20a011

Please choose the answer that best reflects your opinion.

qv20a007 I find using computers to perform medical tasks very alarming

qv20a008 Artificial intelligence makes doctors lazy

qv20a009 Humans have a better overview than computers on what happens in my body

qv20a010 A computer can never compete against the experience of a specialized doctor

qv20a011 I think replacement of doctors by artificial intelligence will happen in the far future

1. Strongly disagree
2. Disagree
3. Neither agree nor disagree
4. Agree
5. Strongly agree

Page 4

if qv20a002 = 1

qv20a012 - qv20a016

Please choose the answer that best reflects your opinion.

qv20a012 Even when computers are used to perform tasks of doctors, humans always remain responsible

qv20a013 I would never blindly trust a computer

qv20a014 Artificial intelligence may prevent errors

qv20a015 Artificial intelligence in medicine will reduce waiting times in healthcare

qv20a016 Artificial intelligence should be used to check human judgement

1. Strongly disagree
2. Disagree
3. Neither agree nor disagree
4. Agree
5. Strongly agree



Page 5

if qv20a002 = 1

qv20a017 - qv20a021

Please choose the answer that best reflects your opinion.

qv20a017 Because of the use of artificial intelligence, fewer doctors are required

qv20a018 When artificial intelligence is used, my personal data may fall into the wrong hands

qv20a019 As a patient I want to be treated as a person, not as a number

qv20a020 I find it worrisome that a computer does not take human feelings into account

qv20a021 When a computer can predict that I will get a disease in the future I want to know that no matter what

1. Strongly disagree
2. Disagree
3. Neither agree nor disagree
4. Agree
5. Strongly agree

Page 6

if qv20a002 = 2

qv20a022 - qv20a025

Please choose the answer that best reflects your opinion.

qv20a022 I find using computers to perform medical tasks a bad idea

qv20a023 I find using computers to perform medical tasks safe

qv20a024 I find using computers to perform medical tasks useful

qv20a025 I find using computers to perform medical tasks nifty

1. Strongly disagree
- 2.
- 3.
- 4.
5. Strongly agree

Page 7

if qv20a002 = 2

qv20a026 - qv20a030

Please choose the answer that best reflects your opinion.

qv20a026 I find using computers to perform medical tasks very alarming

qv20a027 Artificial intelligence makes doctors lazy

qv20a028 Humans have a better overview than computers on what happens in my body

qv20a029 A computer can never compete against the experience of a specialized doctor

qv20a030 I think replacement of doctors by artificial intelligence will happen in the far future

1. Strongly disagree
- 2.
- 3.
- 4.
5. Strongly agree



Page 8

if qv20a002 = 2

qv20a031 - qv20a035

Please choose the answer that best reflects your opinion.

qv20a031 Even when computers are used to perform tasks of doctors, humans always remain responsible

qv20a032 I would never blindly trust a computer

qv20a033 Artificial intelligence may prevent errors

qv20a034 Artificial intelligence in medicine will reduce waiting times in healthcare

qv20a035 Artificial intelligence should be used to check human judgement

1. Strongly disagree
- 2.
- 3.
- 4.
5. Strongly agree

Page 9

if qv20a002 = 2

qv20a036 - qv20a040

Please choose the answer that best reflects your opinion.

qv20a036 Because of the use of artificial intelligence, fewer doctors are required

qv20a037 When artificial intelligence is used, my personal data may fall into the wrong hands

qv20a038 As a patient I want to be treated as a person, not as a number

qv20a039 I find it worrisome that a computer does not take human feelings into account

qv20a040 When a computer can predict that I will get a disease in the future I want to know that no matter what

1. Strongly disagree
- 2.
- 3.
- 4.
5. Strongly agree

Page 10

if qv20a002 = 3

qv20a041

Please choose the answer that best reflects your opinion.

I find using computers to perform medical tasks...

1. A very good idea
2. A good idea
3. Not a bad and not a good idea
4. A bad idea
5. A very bad idea



if qv20a002 = 3

qv20a042

I find using computers to perform medical tasks...

1. Very unsafe
2. Unsafe
3. Neither safe nor unsafe
4. Safe
5. Very safe

if qv20a002 = 3

qv20a043

I find using computers to perform medical tasks...

1. Very useful
2. Useful
3. Not useful and not useless
4. Useless
5. Very useless

if qv20a002 = 3

qv20a044

I find using computers to perform medical tasks...

1. Very convenient
2. Convenient
3. Not convenient and not cumbersome
4. Cumbersome
5. Very cumbersome

Page 11

if qv20a002 = 3

qv20a045

I find using computers to perform medical tasks...

1. Totally not alarming
2. Barely alarming
3. Moderately alarming
4. Reasonably alarming
5. Seriously alarming

if qv20a002 = 3

qv20a046

Artificial intelligence makes doctors...

1. Very active
2. Active
3. Not lazy and not active
4. Somewhat lazy
5. Very lazy



if qv20a002 = 3

qv20a047

Humans have ... overview compared to computers on what happens in my body

1. The best
2. A better
3. Just as well
4. Less
5. Far less

if qv20a002 = 3

qv20a048

A computer can ... compete against the experience of a specialized doctor

1. Never
2. Almost never
3. Sometimes
4. Often
5. Always

if qv20a002 = 3

qv20a049

I think replacement of doctors by artificial intelligence will happen...

1. In the very near future
2. In the near future
3. Not near or far in the future
4. In the far future
5. In the very far future

Page 12

if qv20a002 = 3

qv20a050

Even when computers are used to perform tasks of doctors, humans...

1. Never remain responsible
2. Almost never remain responsible
3. Sometimes remain responsible
4. Often remain responsible
5. Always remain responsible

if qv20a002 = 3

qv20a051

I would ... blindly trust a computer

1. Never
2. Almost never
3. Sometimes
4. Often
5. Always



if qv20a002 = 3

qv20a052

With artificial intelligence, errors can be prevented...

1. Very well
2. Well
3. Not so well and not so badly
4. Badly
5. Very badly

if qv20a002 = 3

qv20a053

Artificial intelligence in medicine will ... waiting times in healthcare

1. Greatly increase
2. Increase
3. Neither increase nor reduce
4. Reduce
5. Greatly reduce

if qv20a002 = 3

qv20a054

Artificial intelligence should ... be used to check human judgement

1. Never
2. Almost never
3. Sometimes
4. Often
5. Always

Page 13

if qv20a002 = 3

qv20a055

Because of the use of artificial intelligence, ...

1. Many more doctors are required
2. More doctors are required
3. An equal number of doctors are required
4. Fewer doctors are required
5. Far fewer doctors are required

if qv20a002 = 3

qv20a056

By using artificial intelligence, the chance of my personal data falling into the wrong hands is...

1. Much smaller
2. Smaller
3. Equal
4. Bigger
5. Much bigger



if qv20a002 =

qv20a057

As a patient I find it ... to be treated as a person and not as a number

1. Very unimportant
2. Unimportant
3. Neither important nor unimportant
4. Important
5. Very important

if qv20a002 = 3

qv20a058

I find it ... that a computer does not take human feelings into account

1. Very worrisome
2. Pretty worrisome
3. Somewhat worrisome
4. Hardly worrisome
5. Not worrisome at all

if qv20a002 = 3

qv20a059

When a computer can predict that I will get a disease in the future, I ... want to know

1. Always
2. Often
3. Sometimes
4. Almost never
5. Never

Page 14

if qv20a002 = 4

qv20a060

Please choose the answer that best reflects your opinion.

I find using computers to perform medical tasks...

1. A very good idea
- 2.
- 3.
- 4.
5. A very bad idea

if qv20a002 = 4

qv20a061

I find using computers to perform medical tasks...

1. Very unsafe
- 2.
- 3.
- 4.
5. Very safe



if qv20a002 = 4

qv20a062

I find using computers to perform medical tasks...

1. Very useful
- 2.
- 3.
- 4.
5. Very useless

if qv20a002 = 4

qv20a063

I find using computers to perform medical tasks...

1. Very convenient
- 2.
- 3.
- 4.
5. Very cumbersome

Page 15

if qv20a002 = 4

qv20a064

I find using computers to perform medical tasks...

1. Totally not alarming
- 2.
- 3.
- 4.
5. Seriously alarming

if qv20a002 = 4

qv20a065

Artificial intelligence makes doctors...

1. Very active
- 2.
- 3.
- 4.
5. Very lazy

if qv20a002 = 4

qv20a066

Humans have ... overview compared to computers on what happens in my body

1. The best
- 2.
- 3.
- 4.
5. Far less



if qv20a002 = 4

qv20a067

A computer can ... compete against the experience of a specialized doctor

1. Never
- 2.
- 3.
- 4.
5. Always

if qv20a002 = 4

qv20a068

I think replacement of doctors by artificial intelligence will happen...

1. In the very near future
- 2.
- 3.
- 4.
5. In the very far future

Page 16

if qv20a002 = 4

qv20a069

Even when computers are used to perform tasks of doctors, humans...

1. Never remain responsible
- 2.
- 3.
- 4.
5. Always remain responsible

if qv20a002 = 4

qv20a070

I would ... blindly trust a computer

1. Never
- 2.
- 3.
- 4.
5. Always

if qv20a002 = 4

qv20a071

With artificial intelligence, erros can be prevented...

1. Very well
- 2.
- 3.
- 4.
5. Very badly



if qv20a002 = 4

qv20a072

Artificial intelligence in medicine will ... waiting times in healthcare

1. Greatly increase
- 2.
- 3.
- 4.
5. Greatly reduce

if qv20a002 = 4

qv20a073

Artificial intelligence should ... be used to check human judgement

1. Never
- 2.
- 3.
- 4.
5. Always

Page 17

if qv20a002 = 4

qv20a074

Because of the use of artificial intelligence, ...

1. Many more doctors are required
- 2.
- 3.
- 4.
5. Far fewer doctors are required

if qv20a002 = 4

qv20a075

By using artificial intelligence, the chance of my personal data falling into the wrong hands is...

1. Much smaller
- 2.
- 3.
- 4.
5. Much bigger

if qv20a002 = 4

qv20a076

As a patient I find it ... to be treated as a person and not as a number

1. Very unimportant
- 2.
- 3.
- 4.
5. Very important



if qv20a002 = 4

qv20a077

I find it ... that a computer does not take human feelings into account

1. Verry worrisome
- 2.
- 3.
- 4.
5. Not worrisome at all

if qv20a002 = 4

qv20a078

When a computer can predict that I will get a disease in the future, I ... want to know

1. Always
- 2.
- 3.
- 4.
5. Never

Page 18

{intro_b2}

Artificial intelligence in radiology

The following questions are about radiology. A radiologist is a medical specialist who looks for a disease, injury or condition through radiation or devices that reveal the tissues and organs of the body. In radiology too, we expect that in the future X-rays and scans can (partly) be assessed independently by a computer (artificial intelligence). We would like to know what you think about the future use of (partly) independently operating computers in assessing X-rays and scans.

Page 19

if qv20a002 = 1

qv20a079 - qv20a083

Please choose the answer that best reflects your opinion.

qv20a079 As far as I am concerned, artificial intelligence can replace doctors in evaluating scans

qv20a080 Even if computers are better in evaluating scans, I still prefer a doctor

qv20a081 I think radiology is not ready for implementing artificial intelligence in evaluating scans

qv20a082 It worries me when computers analyze scans without interference of humans

qv20a083 The sooner I get the results, even when this is from a computer, the more I am at ease

1. Strongly disagree
2. Disagree
3. Neither agree nor disagree
4. Agree
5. Strongly agree



Page 20

if qv20a002 = 1

qv20a084 - qv20a088

Please choose the answer that best reflects your opinion.

qv20a084 Through human experience a radiologist can detect more than the computer

qv20a085 It is unclear to me how computers will be used in evaluating scans

qv20a086 When discussing the results of a scan, talking to a human being is very important

qv20a087 When a computer gives the result of a scan, I would miss the human explanation

qv20a088 Getting the results of a scan involves personal contact

1. Strongly disagree
2. Disagree
3. Neither agree nor disagree
4. Agree
5. Strongly agree

Page 21

if qv20a002 = 1

qv20a089 - qv20a093

Please choose the answer that best reflects your opinion.

qv20a089 I find it important to be able to ask questions when getting the result of a scan

qv20a090 If a computer would give the results of a scan, I would not feel emotional support

qv20a091 I wonder how it is possible that a computer can give me the results of a scan

qv20a092 A computer should only look at body parts that were selected by my doctor

qv20a093 If it does not matter in costs, a computer should always make a full body scan instead of looking at specific body parts

1. Strongly disagree
2. Disagree
3. Neither agree nor disagree
4. Agree
5. Strongly agree



Page 22

if qv20a002 = 2

qv20a094 - qv20a098

Please choose the answer that best reflects your opinion.

qv20a094 As far as I am concerned, artificial intelligence can replace doctors in evaluating scans

qv20a095 Even if computers are better in evaluating scans, I still prefer a doctor

qv20a096 I think radiology is not ready for implementing artificial intelligence in evaluating scans

qv20a097 It worries me when computers analyze scans without interference of humans

qv20a098 The sooner I get the results, even when this is from a computer, the more I am at ease

1. Strongly disagree
- 2.
- 3.
- 4.
5. Strongly agree

Page 23

if qv20a002 = 2

qv20a099 - qv20a103

Please choose the answer that best reflects your opinion.

qv20a099 Through human experience a radiologist can detect more than the computer

qv20a100 It is unclear to me how computers will be used in evaluating scans

qv20a101 When discussing the results of a scan, talking to a human being is very important

qv20a102 When a computer gives the result of a scan, I would miss the human explanation

qv20a103 Getting the results of a scan involves personal contact

1. Strongly disagree
- 2.
- 3.
- 4.
5. Strongly agree



Page 24

if qv20a002 = 2

qv20a104 - qv20a108

Please choose the answer that best reflects your opinion.

qv20a104 I find it important to be able to ask questions when getting the result of a scan

qv20a105 If a computer would give the results of a scan, I would not feel emotional support

qv20a106 I wonder how it is possible that a computer can give me the results of a scan

qv20a107 A computer should only look at body parts that were selected by my doctor

qv20a108 If it does not matter in costs, a computer should always make a full body scan instead of looking at specific body parts

1. Strongly disagree
- 2.
- 3.
- 4.
5. Strongly agree

Page 25

if qv20a002 = 3

qv20a109

Please choose the answer that best reflects your opinion.

As far as I am concerned, artificial intelligence can ... doctors in evaluating scans

1. Not replace
2. Minimally replace
3. Partially replace
4. Largely replace
5. Completely replace

if qv20a002 = 3

qv20a110

Even if computers are better in evaluating scans, I still ... prefer a doctor

1. Do not
2. Slightly
3. Partially
4. Largely
5. Completely

if qv20a002 = 3

qv20a111

I think radiology ... for implementing artificial intelligence in evaluating scans

1. Is totally not ready
2. Is not ready
3. Is partially ready
4. Is largely ready
5. Is completely ready



if qv20a002 = 3

qv20a112

When computers analyze scans without interference of humans, I find that...

1. Very worrisome
2. Worrisome
3. A little worrisome
4. Not worrisome*
5. Not at all worrisome

**) Note: At question qv20a112, the label of answer category 4 was first displayed incorrectly. This category was first shown as "worrisome". On April 14, 2020 at 2:46 pm, this was adjusted to "not worrisome".*

if qv20a002 = 3

qv20a113

Getting the results quickly, even when this is from a computer, ...

1. Really worries me
2. Worries me
3. Reassures me a bit
4. Reassures me
5. Fully reassures me

Page 26

if qv20a002 = 3

qv20a114

Through human experience a radiologist can ... a scan than a computer

1. Much better detect
2. Better detect
3. Equally well detect
4. Less well detect
5. Far less well detect

if qv20a002 = 3

qv20a115

It is ... to me how computers will be used in evaluating scans

1. Very unclear
2. Unclear
3. Neither clear, nor unclear
4. Clear
5. Very clear

if qv20a002 = 3

qv20a116

When discussing the results of a scan, talking to a human being is...

1. Totally unimportant
2. Unimportant
3. Kind of important
4. Important
5. Very important



if qv20a002 = 3

qv20a117

When a computer gives the result of a scan, I would ... the human explanation

1. Totally not miss
2. Not miss
3. Slightly miss
4. Miss
5. Really miss

if qv20a002 = 3

qv20a118

Getting the results of a scan should ... involve personal contact

1. Never
2. Almost never
3. Sometimes
4. Often
5. Always

Page 27

if qv20a002 = 3

qv20a119

I find it ... to be able to ask questions when getting the result of a scan

1. Totally unimportant
2. Unimportant
3. Kind of important
4. Important
5. Very important

if qv20a002 = 3

qv20a120

If a computer would give the results of a scan, I would...

1. Feel a lot of emotional support
2. Feel emotional support
3. Feel a little emotional support
4. Feel no emotional support
5. Not feel any emotional support at all

if qv20a002 = 3

qv20a121

I ... how it is possible that a computer can give me the results of a scan

1. Do not wonder at all
2. Do not wonder
3. Wonder a little
4. Wonder a lot
5. Wonder very much



if qv20a002 =

qv20a122

Of the body parts selected by a treating doctor, a computer should look at...

1. Body parts that were selected by my doctor only
2. The selection and a limited part of adjacent body parts
3. The selection and a reasonable part of adjacent body parts
4. The selection and all adjacent body parts
5. The entire body

if qv20a002 = 3

qv20a123

If it does not matter in costs, a computer should ... make a full body scan instead of looking at specific body parts

1. Never
2. Almost never
3. Sometimes
4. Often
5. Always

Page 28

if qv20a002 = 4

qv20a124

Please choose the answer that best reflects your opinion.

As far as I am concerned, artificial intelligence can ... doctors in evaluating scans

1. Not replace
- 2.
- 3.
- 4.
5. Completely replace

if qv20a002 = 4

qv20a125

Even if computers are better in evaluating scans, I still ... prefer a doctor

1. Do not
- 2.
- 3.
- 4.
5. Completely

if qv20a002 = 4

qv20a126

I think radiology ... for implementing artificial intelligence in evaluating scans

1. Is totally not ready
- 2.
- 3.
- 4.
5. Is completely ready



if qv20a002 = 4

qv20a127

When computers analyze scans without interference of humans, I find that...

1. Very worrisome
- 2.
- 3.
- 4.
5. Not at all worrisome

if qv20a002 = 4

qv20a128

Getting the results quickly, even when this is from a computer, ...

1. Really worries me
- 2.
- 3.
- 4.
5. Fully reassures me

Page 29

if qv20a002 = 4

qv20a129

Through human experience a radiologist can ... a scan than a computer

1. Much better detect
- 2.
- 3.
- 4.
5. Far less well detect

if qv20a002 = 4

qv20a130

It is ... to me how computers will be used in evaluating scans

1. Very unclear
- 2.
- 3.
- 4.
5. Very clear

if qv20a002 = 4

qv20a131

When discussing the results of a scan, talking to a human being is...

1. Totally unimportant
- 2.
- 3.
- 4.
5. Very important



if qv20a002 = 4

qv20a132

When a computer gives the result of a scan, I would ... the human explanation

1. Totally not miss
- 2.
- 3.
- 4.
5. Really miss

if qv20a002 = 4

qv20a133

Getting the results of a scan should ... involve personal contact

1. Never
- 2.
- 3.
- 4.
5. Always

Page 30

if qv20a002 = 4

qv20a134

I find it ... to be able to ask questions when getting the result of a scan

1. Totally unimportant
- 2.
- 3.
- 4.
5. Very important

if qv20a002 = 4

qv20a135

If a computer would give the results of a scan, I would...

1. Feel a lot of emotional support
- 2.
- 3.
- 4.
5. Not feel any emotional support at all

if qv20a002 = 4

qv20a136

I ... how it is possible that a computer can give me the results of a scan

1. Do not wonder at all
- 2.
- 3.
- 4.
5. Wonder very much



Page 46

Of the body parts selected by a treating doctor, a computer should look at...

1. Body parts that were selected by my doctor only
- 2.
- 3.
- 4.
5. The entire body

if qv20a002 = 4

qv20a138

If it does not matter in costs, a computer should ... make a full body scan instead of looking at specific body parts

1. Never
 - 2.
 - 3.
 - 4.
 5. Always
-

Annex II Enquesta traduïda al català ciutadania

--	--	--



Intel·ligència artificial aplicada a la salut

Des de la Gerència Territorial de la Catalunya Central de l'Institut Català de la Salut s'està realitzant un estudi sobre la percepció de la intel·ligència artificial a la salut. Agrairíem la seva col·laboració responant aquesta enquesta fàcil, ràpida i totalment anònima.

Els avenços en l'àmbit de la tecnologia van molt ràpids. En un futur, molts procediments mèdics podran ser, en part, realitzats de manera independent per un ordinador. Aquesta ajuda dels ordinadors és també anomenada "Intel·ligència Artificial" (IA).

Ens agradaria saber que pensa sobre aquest ús dels ordinadors que funcionen, en part, de manera independent realitzant tasques mèdiques.

Un cop contestada l'enquesta, recordi lliurar-la abans d'abandonar el Centre d'Atenció Primària. Moltes gràcies per contribuir en la recerca.

1. Si us plau, indiqui el seu gènere:

Dona Home Altres

--	--

2. Indiqui la seva edat en anys:

3. Indiqui el seu estat civil:

- Solter/a
- Casat/da
- Divorciat/da
- Vidu/a
- Altres

4. Indiqui el seu nivell d'estudis:

- No sap / No contesta
- Primària
- Secundària
- Batxillerat, FP
- Universitaris

--	--	--	--	--

5. Indiqui el codi postal de la seva població:



9578693

6. Si us plau, triï la resposta que millor reflecteixi la seva opinió:

Totalment en desacord	En desacord	Indiferent	D'acord	Totalment d'acord
1	2	3	4	5

	<u>Totalment en desacord</u>			<u>Totalment d'acord</u>	
Considero que utilitzar els ordinadors per realitzar tasques mèdiques és una mala idea.	<input type="checkbox"/> 1	<input type="checkbox"/> 2	<input type="checkbox"/> 3	<input type="checkbox"/> 4	<input type="checkbox"/> 5
Considero que utilitzar els ordinadors per realitzar tasques mèdiques és segur.	<input type="checkbox"/> 1	<input type="checkbox"/> 2	<input type="checkbox"/> 3	<input type="checkbox"/> 4	<input type="checkbox"/> 5
Considero que utilitzar els ordinadors per realitzar tasques mèdiques és útil.	<input type="checkbox"/> 1	<input type="checkbox"/> 2	<input type="checkbox"/> 3	<input type="checkbox"/> 4	<input type="checkbox"/> 5
Considero que utilitzar els ordinadors per realitzar tasques mèdiques és enginyós.	<input type="checkbox"/> 1	<input type="checkbox"/> 2	<input type="checkbox"/> 3	<input type="checkbox"/> 4	<input type="checkbox"/> 5
Considero que utilitzar els ordinadors per realitzar tasques mèdiques és molt alarmant.	<input type="checkbox"/> 1	<input type="checkbox"/> 2	<input type="checkbox"/> 3	<input type="checkbox"/> 4	<input type="checkbox"/> 5

7. Entens què significa el terme d'Intel·ligència Artificial?

No Si

8. Amb les teves paraules, què significa la Intel·ligència Artificial (IA)?



9578693

9. Si us plau, triï la resposta que millor reflecteixi la seva opinió:

Totalment en desacord	En desacord	Indiferent	D'acord	Totalment d'acord
1	2	3	4	5

La IA fa tornar els metges mandrosos.

 1 2 3 4 5

Els humans tenen una millor visió general del que passa al meu cos que els ordinadors.

 1 2 3 4 5

Un ordinador mai podrà competir contra l'experiència d'un metge especialitzat.

 1 2 3 4 5

Considero que reemplaçar els metges per la IA passarà en un futur llunyà.

 1 2 3 4 5

Encara que els ordinadors siguin usats per realitzar tasques mèdiques, els humans sempre seran els responsables.

 1 2 3 4 5

Mai confiaria cegament en un ordinador.

 1 2 3 4 5

La IA pot prevenir errors.

 1 2 3 4 5

La IA en medicina reduirà els temps d'espera en salut.

 1 2 3 4 5

La IA s'hauria d'utilitzar per comprovar decisions humanes.

 1 2 3 4 5

Amb l'ús de la IA, seran necessaris menys metges.

 1 2 3 4 5

Amb l'ús de la IA, les meves dades personals poden caure en mans equivocades.

 1 2 3 4 5

Com a pacient, vull ser tractat com una persona i no com un número.

 1 2 3 4 5

Considero preocupant que un ordinador no tingui en compte els sentiments humans.

 1 2 3 4 5

Si un ordinador pot predir que tindrà una malaltia en un futur, vull saber-ho siguin quines siguin les conseqüències.

 1 2 3 4 5



10. Si us plau, triï la resposta que millor reflecteixi la seva opinió:

Totalment en desacord	En desacord	Indiferent	D'acord	Totalment d'acord
1	2	3	4	5

- Segons la meua opinió la IA pot reemplaçar els metges en l'avaluació de proves radiològiques. 1 2 3 4 5
- Tot i que els ordinadors siguin millor avaluant proves radiològiques, jo continuo preferint que ho faci un metge. 1 2 3 4 5
- Crec que la radiologia no està preparada per implementar la IA en l'avaluació de proves radiològiques. 1 2 3 4 5
- Em preocupa que els ordinadors avaluïn proves radiològiques sense la supervisió humana. 1 2 3 4 5
- Com més ràpid pugui disposar dels resultats, encara que siguin obtinguts d'un ordinador, millor. 1 2 3 4 5
- Amb l'experiència humana, un radiòleg pot interpretar millor una radiografia que un ordinador. 1 2 3 4 5
- No tinc clar com seran utilitzats els ordinadors per avaluar proves radiològiques. 1 2 3 4 5
- Quan es discuteixen els resultats d'una prova radiològica, parlar-ho amb un humà és molt important. 1 2 3 4 5
- Si un ordinador em donés el resultat d'una prova radiològica, trobaria a faltar l'explicació humana. 1 2 3 4 5
- Donar els resultats d'una prova radiològica implica el contacte personal. 1 2 3 4 5
- Considero important tenir la possibilitat de fer preguntes quan em donen els resultats d'una prova radiològica. 1 2 3 4 5
- Si un ordinador donés els resultats d'una prova radiològica, em mancaria el suport emocional humà. 1 2 3 4 5
- Em pregunto com és possible que un ordinador pugui donar-me els resultats d'una prova radiològica. 1 2 3 4 5
- Un ordinador només hauria de mirar les parts del cos que han estat seleccionades pel meu metge. 1 2 3 4 5
- Si no importa el cost econòmic, l'ordinador sempre hauria de fer una prova radiològica del cos complet, en lloc de mirar parts específiques. 1 2 3 4 5

Annex III Cartell informatiu sobre l'enquesta a la ciutadania

Què n'opines de la **INTEL·LIGÈNCIA ARTIFICIAL** aplicada a la salut?



Escaneja el codi
QR amb el mòbil i
accedeix al formulari.



Ajuda'ns a millorar el servei que
t'oferim responent aquesta
**ENQUESTA RÀPIDA, FÀCIL
I ANÒNIMA!**

Annex IV Enquesta original professionals



Faculty of Health/Discipline of Nursing

1. What is your age?
 2. What is your gender?
 3. What sort of district do you work in?
 - Rural/remote
 - Metropolitan
 - Regional
 4. What health discipline do you work in?
 - Nursing/midwifery
 - Medicine
 - Allied Health
 - Other health
 - non-clinical
 5. What job description do you hold?
 - Health employee
 - Manager
 - Educator
 - Other
 6. How many years experience do you have in your current role?
 7. Do you understand what the term Artificial intelligence means?
 - Yes I think so
 - No I don't think so
 8. In your own words, describe what you understand Artificial Intelligence to mean?
-
9. In your current role are you using Artificial intelligence to assist you to deliver or plan care?
 - Yes
 - No
 - I don't know
 10. What do you see as some of the challenges to incorporating Artificial Intelligence in your workplace?
 - Infrastructure
 - Interoperability with current systems
 - Cost to implement
 - Workforce knowledge and skills
 - Organisational support
 - Interdisciplinary collaboration
 - Clinical governance
 - Research funding

www.scu.edu.au



Faculty of Health/Discipline of Nursing

- I don't know
- Other: Please specify _____

11. What sort of education or training have you had about Artificial Intelligence?

- Formal university qualification
- Workplace training
- Self-initiated online course, webinar, conference etc
- None

12. Would you like to receive education about Artificial Intelligence in healthcare?

- Yes
- No

13. Which topics would you like to know more about?

- General teaching about Artificial Intelligence capabilities
- The application of Artificial Intelligence in healthcare
- The ethics of Artificial Intelligence in healthcare
- Artificial Intelligence for your business
- Training on AI techniques: machine learning, deep learning etc.
- Other: _____

14. The following statements measure your perceptions of Artificial Intelligence at this moment in time. Please rate the following questions:

	Totally Disagree	Disagree	Unsure	Agree	Totally Agree
I believe that the use of AI in my specialty could improve the delivery of patient care	1	2	3	4	5
I believe that the use of AI in my specialty could improve clinical decision making	1	2	3	4	5
I believe that AI can improve population health outcomes	1	2	3	4	5
I believe that AI will change my role as a healthcare professional in the future	1	2	3	4	5
I believe that the introduction of AI will reduce financial cost associated with my role	1	2	3	4	5
I believe that overall healthcare professionals are prepared for the introduction of AI technology	1	2	3	4	5
I believe that one day AI may take over part of my role as a healthcare professional	1	2	3	4	5
I believe that I have been adequately trained to use AI that is specific to my role.	1	2	3	4	5

www.scu.edu.au



Faculty of Health/Discipline of Nursing

I believe there is an ethical framework in place for the use of AI technology in my workplace	1	2	3	4	5
I believe that should AI technology make an error; full responsibility lies with the healthcare professional	1	2	3	4	5

To reference this survey please cite: Shinnars, L., Grace, S., Smith, S., Stephens, A., & Aggar, C. (2022). Exploring healthcare professionals' perceptions of artificial intelligence: Piloting the Shinnars Artificial Intelligence Perception tool. *DIGITAL HEALTH*, 8, DOI: 10.1177/20552076221078110 journals.sagepub.com/home/dhj

Annex V Enquesta traduïda professionals



Intel·ligència artificial en salut

Aquesta enquesta forma part d'un estudi de recerca per poder conèixer les percepcions de la intel·ligència artificial (IA) dels professionals i dels usuaris d'atenció primària.

Tot i ser conceptes recents, el ràpid creixement de la informàtica i el gran augment de dades, indiquen que la intel·ligència artificial ha vingut per quedar-se i alterar de forma significativa la pràctica de la medicina. En aquest moment, és essencial conèixer la percepció en vers la IA, tant dels professionals com dels usuaris, per tal d'establir protocols i guies d'actuació per afavorir l'acceptació d'aquesta tecnologia, així com per superar les preocupacions que es poden derivar del seu ús.

Es tracta d'una enquesta anònima i amb tractament confidencial de dades; no hauràs de dedicar-hi més de 3 minuts a respondre-la.

* Necessària

1. Quina és la teva edat (en anys) *

Escriviu la vostra resposta.

2. Quin és el teu sexe *

Home

Dona

3. En quina disciplina de la salut treballes? *

Infermeria

Medicina

Altres

4. Quants d'anys d'experiència tens al teu lloc de treball actual? *

Menys de 1 any

Entre 1 i 2 anys

Entre 3 i 4 anys

Entre 5 i 10 anys

Més de 10 anys

5. Entens què significa el concepte de intel·ligència artificial? *

Sí

No

6. Descriu breument el que per tu significa la intel·ligència artificial? *

Escriviu la vostra resposta.

7. A la teva feina actual fas servir la intel·ligència artificial per ajudar-te a realitzar o planificar l'atenció? *

Sí

No

No ho sé

8. Quins creus que són els reptes d'incorporar la intel·ligència artificial al teu lloc de treball? *

Infraestructura

Interoperabilitat amb els sistemes actuals

Cost d'implementació

Coneixements i habilitat de la plantilla de professionals

Suport organitzatiu

Col·laboració interdisciplinària

- Governança clínica
- Finançament de la recerca
- No ho sé
- Altres

9. Quin tipus d'educació o formació has tingut sobre la intel·ligència artificial ? *

- Formació universitària
- Formació del lloc de treball
- Cursos en línia d'iniciació pròpia, seminaris web, conferències, etc.
- Cap

10. T'agradaria rebre educació sobre la intel·ligència artificial en salut ? *

- Sí
- No

11. De quins temes t'agradaria saber-ne més ? *

- Ensenyament general sobre les capacitats de la intel·ligència artificial
- L'aplicació de la intel·ligència artificial en sanitat
- L'ètica de la intel·ligència artificial en sanitat
- Intel·ligència artificial pel teu negoci
- Formació en tècniques de intel·ligència artificial: aprenentatge automàtic, aprenentatge profund, etc.
- Altres

12. Les afirmacions següents mesuren les vostres percepcions de la intel·ligència artificial (IA) en aquest moment. Valoreu les preguntes següents: *

	Totalment en desacord	En desacord	Ni d'acord ni en desacord	D'acord	Totalment d'acord
Crec que l'ús de la IA en la meua especialitat podria millorar l'atenció al pacient	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Crec que l'ús de la IA en la meua especialitat podria millorar la presa de decisions clíniques	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Crec que la IA pot millorar els indicadors de salut de la població	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Crec que la IA pot canviar el meu rol com a professional sanitari en un futur	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Crec que la introducció de la IA pot reduir el cost econòmic associat a la meua feina	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Crec que la majoria dels professionals sanitaris estem preparats per la introducció de la tecnologia de la IA	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Crec que un dia la IA podrà fer part de la meua feina com a professional sanitari	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Crec que he estat adequadament ensenyat per utilitzar la IA en el meu lloc de treball	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Crec que hi ha un marc ètic per l'ús de la tecnologia d'IA en el meu lloc de treball	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Crec que si la tecnologia d'IA comet un error, tota la responsabilitat recau sobre el professional sanitari	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

13. A continuació es mostren possibles aplicacions de la IA en radiologia. Selecciona on creus que es podria aplicar i quina prioritat hauria de tenir: *

	Prioritat alta	Prioritat moderada	Prioritat baixa	La IA no s'hauria d'utilitzar en aquesta àrea	Sense opinió
Eina de suport en les decisions clíniques per la correcta justificació de proves complementàries	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Interpretació d'imatges per resultats preliminars immediats	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Interpretació d'imatges per un resultat final i definitiu del pacient	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Envia

Annex VI Llistat de patologies detectades amb l'algorisme d'IA i agrupació

Llistat de patologies detectades a través de l'algorisme d'IA *ChestEye* i la seva agrupació

Patologia	Agrupació
Adhesió/adherència pleural	Patologia pleural
Aorta augmentada	Patologia de vasos
Atelèctasi laminar	Patologia parènquima pulmonar
Barra Nuss/Pectus excavatum	Patologia òssia
Cables externals	Implants externs
Calcificació gangli limfàtic	Altres
Canvis degeneratius vertebrals/artrosis	Patologia òssia
Cardiomegàlia	Patologia cardíaca i/o valvular
Cifosi	Patologia òssia
Col·locació de catèter	Implants externs
Congestió	Patologia parènquima pulmonar
Consolidació	Patologia parènquima pulmonar
Costella anormal	Patologia òssia
Desplaçament mediastínic	Mediastí
Eixamplament hilar o prominència hilar	Patologia parènquima pulmonar
Elevació diafragma	Patologia abdomen superior
Emfisema bullós	Patologia parènquima pulmonar
Emfisema pulmonar/ atrapament aeri	Patologia parènquima pulmonar
Engruiximent de la cissura	Patologia parènquima pulmonar
Engruiximent pleural	Patologia pleural
Entesopatia vertebral	Patologia òssia
Esclerosi aòrtica	Patologia de vasos
Escoliosi	Patologia òssia
Fibrosi pulmonar	Patologia parènquima pulmonar
Fractura vertebral	Patologia òssia
Gasos intestinals	Patologia abdomen superior
Granuloma	Altres
Hèrnia	Patologia abdomen superior
Hipertensió pulmonar	Patologia de vasos
Hipoventilació	Patologia parènquima pulmonar
Implant espinal	Patologia òssia
Limfadenopatia	Altres
Marcapassos	Implants externs
Marques intersticials	Patologia parènquima pulmonar
Massa	Patologia parènquima pulmonar
Mediastí eixamplat	Mediastí
Nòdul	Patologia parènquima pulmonar
Pneumomediastí	Mediastí
Pneumoperitoneu	Patologia abdomen superior
Pneumotòrax	Patologia pleural
Sarcoïdosi	Patologia parènquima pulmonar
Tuberculosi	Patologia parènquima pulmonar

Vàlvula cardíaca artificial	Patologia cardíaca i/o valvular
Vessament pleural	Patologia pleural