

ORIGINAL

Experiencia de Mediktor®: un nuevo evaluador de síntomas basado en inteligencia artificial para pacientes atendidos en el servicio de urgencias

Elvira Moreno Barriga¹, Irene Pueyo Ferrer², Miquel Sánchez Sánchez², Montserrat Martín Baranera³, Josep Masip Utset¹

Objetivo. Analizar la concordancia entre los diagnósticos emitidos por Mediktor® con el realizado por el médico responsable, así como valorar la utilidad de este dispositivo en pacientes que acuden a un servicio de urgencias (SU).

Método. Estudio observacional prospectivo realizado en el SU de un hospital terciario universitario. A los pacientes con patologías médicas y quirúrgicas (cirugía y traumatología) que no precisaban asistencia médica inmediata se les entregó una tableta digital para responder al interrogatorio de Mediktor®. Según las respuestas, el software adjudicaba un listado de 10 prediagnósticos ordenados por probabilidad, que se ocultaban al paciente y al médico responsable, para no modificar el proceso habitual. Posteriormente se analizó el grado de coincidencia entre el diagnóstico médico y los diagnósticos ofrecidos por Mediktor®.

Resultados. 1.015 pacientes fueron incluidos, de los que 622 se consideraron casos válidos para el estudio. Se excluyeron los pacientes que no cumplían los criterios de inclusión, sin diagnóstico al alta, con diagnóstico final expresado como síntoma y aquellos con diagnósticos no incluidos en Mediktor®. Las coincidencias entre el diagnóstico médico (patrón oro) y los diez diagnósticos de Mediktor® fueron de un 91,3%, en los tres primeros diagnósticos de un 75,4% y en el primer diagnóstico de un 42,9%. Según los grupos de diagnósticos más frecuentes, se objetivó una sensibilidad > 92% y una especificidad > 91% en la mayoría de ellos, con un índice kappa que osciló entre el 0,24 y el 0,98.

Conclusiones. Mediktor® es una herramienta fiable para ayudar al diagnóstico de las enfermedades más prevalentes de un SU y fácil de utilizar por el público en general.

Palabras clave: Servicio de Urgencias. Concordancia. Cirugía. Traumatología. Síntoma. Diagnóstico médico.

A new artificial intelligence tool for assessing symptoms in patients seeking emergency department care: the Mediktor application

Objectives. To analyze agreement between diagnoses issued by the Mediktor application and those of an attending physician, and to evaluate the usefulness of this application in patients who seek emergency care.

Methods. Prospective observational study in a tertiary care university hospital emergency department. Patients with medical problems and surgical conditions (surgery and injuries) who did not require immediate emergency care responded to the Mediktor questions on a portable computer tablet. The software analyzed the answers and provided a list of 10 possible preliminary diagnoses in order of likelihood. The patient and the attending physician were blinded to the list so that the usual care process would not be altered. The level of agreement between the physician's diagnosis and the Mediktor diagnosis was analyzed.

Results. A total of 1015 patients were included; 622 cases were considered valid for study. Cases were excluded if the patients did not meet the inclusion criteria, they did not have a discharge diagnosis, they had a final diagnosis expressed as a symptom or their final diagnosis was not included in the Mediktor database. The physician's diagnosis (the gold standard) coincided with one of the 10 MEDIKTOR diagnoses in 91.3% of the cases, with one of the first 3 diagnoses in 75.4%, and with the first diagnosis in 42.9%. Sensitivity was over 92% and specificity over 91% in the majority of common diagnostic groups; the κ statistic ranged from 0.24 to 0.98.

Conclusions. The Mediktor application is a reliable diagnostic aid for the most prevalent problems treated in a hospital emergency department. The general public finds it easy to use.

Keywords: Emergency department. Agreement. Surgery. Injuries. Symptoms. Physician diagnosis.

Introducción

La proliferación de internet y la fácil accesibilidad a la información médica han provocado que actualmente la búsqueda en la red de términos médicos se haya

convertido en una práctica habitual en la población general¹. Si esta información no está bien contrastada puede conducir a interpretaciones erróneas² y, por lo tanto, comportar un riesgo para la salud del usuario. En la línea de mejorar la desinformación existente, han

Filiación de los autores:

¹Servicio de Medicina Intensiva del Consorci Sanitari Integral, Barcelona, España.

²Área de Urgencias, Hospital Clínic, Barcelona, España.

³Departamento de Epidemiología e Investigación del Consorci Sanitari Integral, Barcelona, España.

Contribución de los autores:

Todos los autores han confirmado su autoría en el documento de responsabilidades del autor, acuerdo de publicación y cesión de derechos a EMERGENCIAS.

Autor para correspondencia:

Elvira Moreno Barriga
C/ Ricart, 16, 2º, 1º
08004 Barcelona, España.

Correo electrónico:

elviramorenobarriga@hotmail.com

Información del artículo:

Recibido: 4-5-2017

Aceptado: 7-8-2017

Online: 9-10-2017

Editor responsable:

Agustín Julián-Jiménez, MD, PhD.

proliferado sofisticados programas informáticos denominados evaluadores de síntomas (*symptomcheckers*)³⁻⁵ que, aplicando algoritmos computarizados, interrogan al usuario sobre sus síntomas, para finalmente ofrecer un listado de posibles diagnósticos, indicar con qué urgencia deben solicitar ayuda médica y a quién deben dirigirse. El verdadero impacto del uso de estos evaluadores de síntomas dependerá en gran medida de su grado de precisión y del beneficio clínico y económico que ofrezcan^{6,7}.

En esta línea, MEDIKTOR CORP (M.C) - TECKEL SOLUTION es una empresa cuyo objetivo es desarrollar y comercializar soluciones informáticas basadas en inteligencia artificial centradas en el ámbito médico. Ha desarrollado un evaluador de síntomas denominado Mediktor®, destinado a facilitar y agilizar el proceso de diagnóstico médico mediante algoritmos de inteligencia artificial. A diferencia de otros evaluadores de síntomas que existen en el mercado³⁻⁵, y debido a que la precisión diagnóstica puede verse afectada por la interacción entre el usuario y el *software*, M.C. desarrolla y aplica en Mediktor® algoritmos de interpretación del lenguaje natural, sin requerir el conocimiento de términos técnicos por parte del usuario.

Para la mejora en el proceso de diagnóstico, los estudios que se habían realizado sobre algunos de los distintos evaluadores de síntomas que existían en el mercado se hicieron midiendo su precisión a través de "viñetas de pacientes virtuales". De esta forma obtenían como promedio un 58% de concordancia entre diagnósticos, teniendo en cuenta los 20 primeros diagnósticos otorgados⁸.

Por todo ello, el objetivo de este estudio fue analizar, en una serie de pacientes reales, cual es la concordancia entre los diagnósticos emitidos por Mediktor® tras completar el interrogatorio guiado (realiza un diagnóstico diferencial con diez posibles diagnósticos desde mayor probabilidad a menor) con el diagnóstico emitido por el médico responsable (patrón oro) para el mismo paciente en el mismo episodio y, de esta manera, valorar su utilidad en pacientes que acuden a un servicio de urgencias (SU).

Método

Estudio observacional prospectivo realizado en el SU del Hospital Clínic de Barcelona (HCB). Se trata de un hospital de 800 camas, localizado en el centro de la ciudad, de tercer nivel (alta complejidad) y de referencia de unos 550.000 habitantes. En la actualidad, se atienden anualmente unas 95.000 urgencias generales, excepto de las especialidades de obstetricia y ginecología, pediatría y oftalmología, que se realizan en otros centros. Desde febrero del 2009 el triaje de pacientes en el SU lo realiza enfermería, utilizando la asignación del nivel de urgencia según el Modelo Andorrano de Triage (MAT)⁹. El estudio se realizó en un período de 4 meses, desde el día 1 de octubre 2016 hasta el 31 de enero 2017.

Los pacientes seleccionados fueron aquellos a los que el triaje asignó niveles de MAT III, IV y V (que no requerían atención inmediata) según el orden de llegada y en turnos diferentes para asegurar la obtención de una muestra representativa. Los criterios de inclusión fueron: edad > 18 años, adecuado nivel de consciencia y capacidad de entender la lengua española para poder contestar a las preguntas del programa. Se excluyeron los pacientes que no cumplían los criterios de inclusión, los que denegaron participar en el estudio, aquellos que no tuvieron diagnóstico al alta ("USUARIOS"), aquellos en los que el diagnóstico final fuera expresado como un síntoma o signo ("INESPECÍFICO" = diagnósticos no admitidos en CIE-10) y los casos con diagnósticos no incluidos en el diccionario de Mediktor® en el momento de la realización del estudio ("NO CONTEMPLADO").

Una enfermera adscrita al HCB y destinada al estudio (EE) entrevistó a los pacientes en las plantas de urgencias de medicina, cirugía y traumatología, tras haber sido triados (niveles III-V) y distribuidos siguiendo el proceso habitual. Se presentó mientras los pacientes aguardaban en la sala de espera a ser visitados y les explicó en qué consistía el estudio en un lenguaje entendible (sin tecnicismos), asegurando que el estudio no modificaría el tiempo, las pruebas diagnósticas ni el tratamiento que le pudieran aplicar por su dolencia. Tras confirmar los criterios de inclusión y descartar los criterios de exclusión, se estableció un tiempo prudente de reflexión, tras el cual se solicitó el consentimiento informado por escrito para su inclusión en el estudio y se entregó una copia del mismo. A los pacientes que accedían a participar en el estudio, el programa les asignó un código aleatorizado de 5 números para poder identificarlos. Cada uno de los códigos fue generado por el ordenador, siendo los datos del usuario completamente anónimos puesto que no incluían iniciales, fecha de nacimiento, ni número de documento nacional de identidad. La EE registraba este código en un archivo protegido con el número de HC (historia clínica) del paciente. Este proceso aseguró la codificación aleatoria y el anonimato del participante. Tras la asignación del código, la EE comprobaba los signos vitales del paciente (presión arterial, frecuencia cardíaca, temperatura y pulsioximetría) rellenando los apartados correspondientes en Mediktor®. Posteriormente, les entregaba la tableta digital para que introdujesen el motivo de consulta por el cual habían acudido al SU y respondiesen ellos mismos a las preguntas que el programa les fuese planteando (entre 7 y 20), realizando así la evaluación de manera individual (autoadministrada). El número medio de preguntas que el evaluador realizó entre todos los pacientes incluidos en el estudio fue de 17 preguntas y el tiempo medio calculado por cada sesión fue de 2 minutos y 25 segundos.

Algunas de las preguntas se respondían al ser seleccionadas, otras tenían un formato multirresposta (texto, imagen o la combinación de ambas), y la gran mayoría se respondían escogiendo entre cinco posibles respuestas: SÍ, NO, PROBABLEMENTE SÍ, PROBABLEMENTE NO, o NO LO SÉ. Al finalizar el interrogatorio, el *software* agradecía su participación y les deseaba una

pronta recuperación sin mostrar el listado de diagnósticos, los cuales quedaron registrados de forma oculta en la tableta digital. De esta manera se evitó influenciar al médico responsable, o generar conflictos si hubieran existido diferencias diagnósticas entre el médico y Mediktor®. Al finalizar el cuestionario se invitaba a los pacientes a responder cuatro preguntas sobre la utilidad de Mediktor®. Las preguntas fueron las siguientes: 1) ¿ha entendido bien la mayoría de las preguntas?, 2) ¿la aplicación le ha parecido fácil de usar?, 3) ¿cree que las nuevas tecnologías pueden ayudar a mejorar el sistema sanitario?, y 4) ¿le gustaría poder obtener una orientación médica desde su casa? Todas ellas se respondían con NO/SÍ/NS/NC.

El protocolo del estudio fue aprobado por el Comité de Ética en Investigación Clínica del Hospital Clínic de Barcelona.

El análisis estadístico, el procesamiento de los datos y el análisis de la concordancia se realizaron de forma diferida, revisando la HC del paciente y los diagnósticos registrados por el programa, introduciéndolos en una base de datos. Se establecieron campos diagnósticos de similitud que permitían agrupar diagnósticos expresados con distinta terminología.

Dado que no existían antecedentes en la literatura, se realizó un estudio piloto con 358 pacientes para estimar el tamaño de la muestra. Tras el análisis estadístico y teniendo en cuenta el porcentaje de pérdidas asociado, se estimó como necesario un mínimo de 588 casos. Los datos se analizaron con el programa estadístico SPSS v20 (IBM®, North Castle, Nueva York, EE.UU.), con el cual se realizaron varios cálculos:

- La proporción de coincidencias entre Mediktor® y el médico responsable como el número de coincidencias sobre el total de casos válidos. Se tuvieron en cuenta las coincidencias en el primer diagnóstico, en los tres primeros diagnósticos y en los diez primeros diagnósticos proporcionados por Mediktor®. Para cada una de las tres proporciones, se calcularon los intervalos de confianza (IC) al 95% mediante el método exacto de Clopper-Pearson.

- Teniendo en cuenta los diagnósticos más frecuentes en el SU y considerando el diagnóstico médico al alta como patrón de referencia, se calculó la sensibilidad (S) (probabilidad de que Mediktor® incluyera entre sus diez primeros diagnósticos el diagnóstico clínico considerado, cuando la enfermedad estaba presente), especificidad (E) (probabilidad de que el resultado de Mediktor® fuera negativo cuando la enfermedad no estaba presente), valor predictivo positivo (VPP) (probabilidad de que la enfermedad estuviera presente cuando Mediktor® la incluía entre sus diez primeros diagnósticos) y valor predictivo negativo (VPN) (probabilidad de que la enfermedad no estuviera presente cuando Mediktor® ofrecía un resultado negativo para esa enfermedad), expresándose cada uno de ellos como porcentajes. Los IC al 95% para la sensibilidad y la especificidad se calcularon mediante la fórmula exacta de Clopper-Pearson. Para los IC al 95% de los valores predictivos se aplicó el método logit estándar de Mercado *et al.*¹⁰. La metodología de valoración de fiabili-

dad se realizó mediante el cálculo del índice kappa con su IC específico para cada categoría diagnóstica.

Resultados

Durante el periodo de estudio se seleccionaron 1.015 pacientes según los criterios de inclusión considerándose definitivamente como casos válidos para el análisis un total de 622 pacientes. Se excluyeron del estudio 393 pacientes, entre los que se encontraban aquellos que no tenían diagnóstico al alta (“USUARIO”, 61 casos), los menores de 18 años (“NO 18”, 21 casos), aquellos que el diagnóstico final fue expresado como un síntoma (“INESPECÍFICO” = diagnósticos no admitidos en CIE-10, 151 casos) y los pacientes con diagnósticos no incluidos en aquel momento en el diccionario de Mediktor® (“NO CONTEMPLADO”, 160 casos).

La variabilidad de la muestra se representa atendiendo al sexo (mujeres: 56,0% y hombres: 44,0%), a las franjas de edad expresadas en años (de 18-24: 26,7%; 25-49: 52,9%; 50-75: 18,3% y mayores de 75: 2,1%), así como a los turnos de trabajo en los que se recogió el total de los casos (de 8-14 horas: 53,7% y de 14-20 horas: 46,3%).

Los diagnósticos más prevalentes dentro del SU, teniendo como patrón de referencia el diagnóstico médico al alta de urgencias, fueron definidos como aquellos grupos de patologías en los que había más de 4 casos recogidos en nuestra serie. Se puede observar en cada uno de ellos el grado de acuerdo del evaluador en los diez diagnósticos otorgados por Mediktor®, en los tres primeros diagnósticos y en el primer diagnóstico (Tabla 1).

Las coincidencias, analizando el global de pacientes incluidos (622), entre el diagnóstico del médico y los diez diagnósticos de Mediktor® fueron en 568 casos, 91% (IC 95%: 88% a 93%); en los tres primeros diagnósticos en 469, 75% (IC 95%: 71% a 78%); y en el primer diagnóstico en 267 pacientes, 43% (IC 95%: 39% a 46%) (Tabla 2).

Los resultados en cuanto a la S, E, VPN y VPP se presentan para los doce diagnósticos más frecuentes (Tabla 3), objetivándose una $S > 92\%$ en 9 de los 12 grupos analizados y una $E > 91\%$ en 10 de ellos. Asimismo, el índice kappa calculado para los mismos grupos diagnósticos osciló entre el 0,24 y el 0,98, obteniéndose el valor más bajo para el grupo de afectación del tracto urinario y el más alto para el grupo de infección de vías respiratorias superiores (Tabla 4).

El grado de utilidad del evaluador se valoró con las 4 preguntas que Mediktor® realizaba al final del cuestionario. Estas y los resultados a las mismas se muestran en la Tabla 5. Para más de un 98 % de los encuestados Mediktor® es un programa fácil de utilizar.

Discusión

Con el fin de encontrar una nueva herramienta de ayuda al diagnóstico clínico decidimos realizar este estudio sobre Mediktor®, ya que, además de ser un eva-

Tabla 1. Diagnósticos más prevalentes en el servicio de Urgencias y coincidencias entre el diagnóstico médico y Mediktor®

Diagnósticos en el servicio de urgencias	Coincidencias entre el diagnóstico médico y Mediktor®		
	Como primer diagnóstico (%)	En los 3 primeros diagnósticos (%)	En los 10 primeros diagnósticos (%)
	Infección respiratoria de vías superiores*	76,2	93,7
Afectación del tracto urinario	70,6	87,2	100
Lumbociatalgia	82,2	93,3	100
Otitis	73,3	96,6	100
Contusión	76	88	100
Esguince*	79,2	87,5	91,7
Dolor abdominal inespecífico	4,5	45,4	100
Cefalea	88,9	88,88	100
Tendinitis*	81,2	81,2	81,2
Gastroenteritis aguda	80	100	100
Vértigo*	84,6	84,6	84,6
Cervicalgia	53,8	61,5	100
Otros diagnósticos**			

*Se consigue menos del 100% de concordancia sumando los casos de las tres columnas.

**Engloba los más de 200 diagnósticos diferentes donde se obtuvieron menos de 4 pacientes incluidos/patología: fisura anal, estreñimiento, panadizo ungueal, herpes zóster, diverticulitis aguda, hemorroides, fascitis plantar, dolor costal mecánico, diarrea del viajero, infección respiratoria de vías bajas, reacción alérgica, etc.

luador de síntomas que utilizaba algoritmos de inteligencia artificial, nos permitía realizar el cuestionario médico sin problemas de comprensión por parte de los usuarios al utilizar un lenguaje natural.

Tras realizar la comparación entre los diagnósticos obtenidos (médico y evaluador), se objetivó que Mediktor® tenía una coincidencia diagnóstica del 91,3% cuando se analizaban los resultados sobre el listado de las 10 posibles patologías señaladas como diagnósticos por Mediktor® (ordenadas por probabilidad), del 75,4% contando los 3 primeros posibles diagnósticos y del 42,9% con el primer diagnóstico, con un grado de acuerdo moderado en la mayoría de pacientes analizados, con un índice kappa que oscilaba entre el 0,24 y el 0,98. A su vez se pudo comprobar la alta sensibilidad (> 92%) y especificidad (> 91%) del evaluador para la mayoría de pacientes analizados.

Tabla 3. Diagnósticos más frecuentes en el servicio de urgencias: sensibilidad, especificidad, valor predictivo positivo y valor predictivo negativo conseguidos por Mediktor®

	Sensibilidad (IC 95) %	Especificidad (IC 95) %	VPP (IC 95) %	VPN (IC 95) %
Infección vías respiratorias superiores	97,5 (91,3-99,7)	100,0 (99,3-100,0)	100	99,6 (98,6-99,9)
Afectación del tracto urinario	92,6 (82,1-97,9)	68,5 (64,5-72,3)	21,8 (19,5-24,4)	99,00 (97,4-99,6)
Lumbociatalgia	94,9 (82,7-99,4)	93,5 (91,2-95,3)	49,3 (41,5-57,2)	99,6 (98,6-99,9)
Otitis	93,7 (79,2-99,2)	93,7 (91,5-95,5)	44,8 (36,9-52,9)	99,6 (98,6-99,9)
Dolor abdominal inespecífico	88,0 (68,8-97,4)	98,0 (96,5-98,9)	64,7 (50,7-76,6)	99,5 (98,5-99,8)
Esguince	100,0 (84,6-100,0)	98,3 (96,9-99,2)	68,7 (54,3-80,3)	100
Contusión	100,0 (86,3-100,0)	93,1 (90,8-95,0)	37,9 (31,2-45,0)	100
Tendinitis	81,2 (54,3-95,9)	93,4 (91,1-95,2)	24,5 (18,2-32,2)	99,5 (98,5-99,8)
Cefalea	100,0 (80,5-100,0)	91,9 (89,4-93,9)	25,8 (20,9-31,2)	100
Gastroenteritis aguda	94,1 (71,3-99,8)	89,3 (86,5-91,6)	19,7 (15,9-24,2)	99,8 (98,8-99,9)
Cervicalgia	100,0 (71,5-100,0)	95,2 (93,2-96,8)	27,5 (21,0-35,1)	100
Vértigo	84,6 (54,5-98,1)	98,7 (97,4-99,4)	57,9 (39,9-73,9)	99,7 (98,8-99,9)

VPP: valor predictivo positivo; VPN: valor predictivo negativo; IC: intervalo de confianza.

Tabla 2. Coincidencias entre el diagnóstico médico y Mediktor® en el total de pacientes (622)

	IC 95%		
	Acierto Mediktor® (%)	Límite inferior	Límite superior
Como primer diagnóstico	267 (42,9)	39,0	46,9
En los 3 primeros diagnósticos	469 (75,4)	71,8	78,7
En los 10 primeros diagnósticos	568 (91,3)	88,8	93,4

Por otra parte, durante la realización del estudio los pacientes se encontraron confiados con el sistema y no tuvieron dificultades a la hora de contestar las diferentes preguntas que el evaluador les iba realizando. Así, más de un 98% de los encuestados consideró que Mediktor® era un programa fácil de utilizar.

El diagnóstico médico de las enfermedades que padecen nuestros pacientes sigue planteándonos retos difíciles, y a pesar de que son escasos los estudios sobre precisión diagnóstica, se sabe que el porcentaje de error en el diagnóstico médico (incluyendo la realización de pruebas complementarias) se sitúa entre el 10 y el 15%¹¹. Llegar a un diagnóstico correcto es un proceso complicado que puede estar influenciado de manera positiva por la facilidad de la memoria para recuperar patrones representativos de enfermedades¹²⁻¹⁴. Aunque, esta capacidad es limitada, ya que no pueden manejarse más de siete posibilidades a la vez y esta está sometida a la posibilidad de sesgos en el razonamiento clínico. Y también, y de manera negativa, por errores vinculados al proceso cognitivo del razonamiento y, asimismo, a factores humanos como el cansancio, la sobrecarga laboral y la escasa organización del tiempo de asistencia entre otros^{15,16}. Hasta el momento, ninguna estrategia, cognitiva¹⁷ o sistémica¹⁸, ha ayudado a mejorar la precisión diagnóstica y, por tanto, a disminuir el error de diagnóstico.

Por tanto, consideramos que los resultados de nuestro estudio aportan datos relevantes que nos llevan a considerar que Mediktor® es una herramienta informática de ayuda al diagnóstico que puede ser complementaria a la práctica clínica habitual, y que su implementación en los SU hospitalarios podría acelerar el proceso

Tabla 4. Diagnósticos más frecuentes en el servicio de urgencias: concordancia (índice kappa) en el diagnóstico entre Mediktor® y el médico de urgencias

Diagnósticos más frecuentes	Índice Kappa	IC 95 (%)
Infección vías respiratorias superiores	0,986	0,965-1,000
Afectación del tracto urinario	0,248	0,185-0,310
Lumbociatalgia	0,618	0,511-0,724
Otitis	0,577	0,459-0,694
Dolor abdominal inespecífico	0,733	0,604-0,862
Esguince	0,807	0,690-0,923
Contusión	0,522	0,397-0,646
Tendinitis	0,351	0,207-0,496
Cefalea	0,383	0,252-0,513
Gastroenteritis aguda	0,295	0,180-0,409
Cervicalgia	0,415	0,247-0,584
Vértigo	0,680	0,492-0,867

IC: intervalo de confianza.

diagnóstico con una mejora de la eficiencia.

Además, y tras el análisis de los resultados, se plantean futuras estrategias con respecto al evaluador, ya que Mediktor® podría convertirse en un complemento al triaje convencional, como triaje avanzado de enfermería¹⁹. De esta manera, podría ser una herramienta de ayuda para adelantar la solicitud de pruebas diagnósticas o administrar tratamientos y, por ende, acortar la estancia en los SU¹⁸. A su vez, se podría plantear si Mediktor® pudiera convertirse en un nuevo sistema de triaje, para lo cual convendría demostrar concordancia con el nivel de urgencia frente al resultado de sistemas validados en la actualidad. Otros algoritmos de inteligencia artificial referenciados lo han intentado sin éxito⁸.

Nuestro estudio presenta una serie de limitaciones. La primera e importante es que se ha realizado en un único centro. Además, los pacientes incluidos solo pertenecían a los grupos III-V del MAT (sistema de triaje) y, por lo tanto, que no requerían atención inmediata por su gravedad. Por otro lado, el porcentaje de pacientes mayores de 75 años fue muy escaso, sin que pueda permitir sacar conclusiones en este grupo etario. En resumen, se puede decir que solo se ha valorado el Mediktor® en el perfil de pacientes jóvenes de bajo riesgo, en los que sí se pueden interpretar los resultados. Por otro lado, un numeroso grupo de pacientes fueron excluidos al ser catalogados como "inespecíficos" por falta de diagnóstico incluido en CIE-10 al alta (14%) o "no contemplado" por no tener el diagnóstico definido en Mediktor® (15%), que en el momento del estudio contemplaba 560 diagnósticos patológicos. En todo caso, ningún estudio hasta la fecha había analizado la concordancia entre diagnósticos médicos y evaluadores de síntomas con pacientes reales.

Tabla 5. Grado de aceptación y utilidad de Mediktor® percibida por los pacientes

Preguntas	Contestación			Total (%)
	Sí (%)	No (%)	NS/NC (%)	
1) ¿Ha entendido bien la mayoría de las preguntas?	98,8	0,7	0,4	100
2) ¿La aplicación le ha parecido fácil de usar?	98,2	1,1	0,7	100
3) ¿Cree que las nuevas tecnologías pueden ayudar a mejorar el sistema sanitario?	95,3	3,9	0,9	100
4) ¿Le gustaría poder obtener una orientación médica desde su casa?	85,3%	13,8	0,9	100

NS/NC: no sabe o no contesta a la pregunta.

Como conclusión, podemos decir que nuestros hallazgos sugieren que Mediktor® tiene una alta coincidencia con el diagnóstico médico, siendo un evaluador sensible y específico, lo que lo convierte en una herramienta útil para ayudar al diagnóstico de las enfermedades más prevalentes de un SU y fácil de utilizar por los pacientes en general. Como futura línea de trabajo se podría plantear la realización de un estudio basado en analizar la concordancia del nivel de urgencia otorgado por Mediktor® frente al resultado de sistemas de triaje ya validados en la actualidad.

Conflicto de intereses

Elvira Moreno Barriga declara que tras la finalización del presente estudio colabora con la empresa MEDIKTOR CORP (M.C) - TECKEL SOLUTIONS que está desarrollando el diccionario médico del evaluador.

El resto de autores declaran no tener ningún conflicto de intereses en relación al presente artículo.

Financiación

Los autores declaran la no existencia de financiación externa del presente artículo.

Responsabilidades éticas

El estudio fue aprobado por el Comité de Ética e Investigación Clínica del Hospital Clínic de Barcelona.

Se obtuvo consentimiento informado de los participantes.

Todos los autores han confirmado el mantenimiento de la confidencialidad y respeto de los derechos de los pacientes en el documento de responsabilidades del autor, acuerdo de publicación y cesión de derechos a EMERGENCIAS.

Artículo no encargado por el Comité Editorial y con revisión externa por pares

Bibliografía

- Barnett ML, Linder JA. Antibiotic prescribing to adults with sore throat in the United States, 1997-2010. *JAMA Intern Med.* 2014;174:138-40.
- Black P. The dangers of using Google as a diagnostic aid. *Br J Nurs.* 2009;18:1157.
- North F, Ward WJ, Varkey P, Tullidge-Scheitel SM. Should you search the Internet for information about your acute symptom? *Telemed J E Health.* 2012;18:213-8.
- Luger TM, Houston TK, Suls J. Older adult experience of online diagnosis: results from a scenario-based think-aloud protocol. *J Med Internet Res.* 2014;16:e16.
- North F, Varkey P, Laing B, Cha SS, Tullidge-Scheitel S. Are e-health web users looking for different symptom information than callers to triage centers? *Telemed J E Health.* 2011;17:19-24.

- 6 Fox S, Duggan M. Health Online 2013. Internet and American life project. Pew Research Center and California Health Care Foundation. (Consultado 7 Agosto 2017). Disponible en: http://www.pewinternet.org/files/old-media/Files/Reports/PIP_HealthOnline.pdf
- 7 Saczynski J, Yarzebski J, Lessard D, Spencer FA, Gurwitz JH, Gore JM, et al. Trends in pre-hospital delay in patients with acute myocardial infarction (from the Worcester Heart Attack Study). *Am J Cardiol.* 2008;102:1589-94.
- 8 Semigran HL, Linder JA, Gidengil C, Mehrotra A. Evaluation of symptom checkers for self diagnosis and triage: audit study. *BMJ.* 2015;351:3480.
- 9 Jiménez JG. Clasificación de pacientes en los servicios de urgencias y emergencias: hacia un modelo de triaje estructurado de urgencias y emergencias. *Emergencias.* 2003;15:165-74.
- 10 Mercado ND, Lau KF, Zhou XH. Confidence intervals for predictive values with an emphasis on case-control studies. *Stat Med.* 2007;26:2170-83.
- 11 Graber ML. The incidence of diagnostic error in Medicine. *BMJ Qual Saf.* 2013;22(Supl 2):ii21-ii27.
- 12 Alonso Carrión L, Castells Bescós E, Alba Conejo E, Ruiz Cantero A. El razonamiento clínico como fuente de error. *Med Clin (Barc).* 2002;118:587-9.
- 13 Alonso Carrión L, Bermúdez F, Abarca M, García Almeida JM. ¿Diagnosticar mejor? Savia nueva para un árbol viejo. *Med Clin (Barc).* 2011;137:561-4.
- 14 Charlin B, Boshuizen HP, Custers EJ, Feltovich PJ. Scripts and clinical reasoning. *Med Educ.* 2007;41:1178-84.
- 15 Scott I. Errors in clinical reasoning: causes and remedial strategies. *BMJ.* 2009;339:22-5.
- 16 Croskerry P. The importance of cognitive errors in diagnosis and strategies to minimize them. *Acad Med.* 2003;78:775-80.
- 17 Graber ML, Kissam S, Payne VL, Meyer AND, Sorensen A, Lenfestey N. Cognitive interventions to reduce diagnostic error: a narrative review. *BMJ Qual Saf.* 2012;21:535-57.
- 18 Singh H, Graber ML, Kissam SM, Sorensen AV, Lenfestey NF, Tant EM, et al. System-related interventions to reduce diagnostic errors: a narrative review. *BMJ Qual Saf.* 2012;21:160-70.
- 19 Carballo Cardona C. Triaje avanzado: es la hora de dar un paso adelante. *Emergencias.* 2015;27:332-5.

FE DE ERRORES

En el artículo "Análisis de redes de colaboración españolas en la investigación en Medicina de Urgencias y Emergencias (2010-2014)" de Burbano Santos *et al.* (*Emergencias.* 2017;29:320-6), la Figura 2 se imprimió de forma demasiado tenue y algunas de las líneas conectoras no se visualizaron correctamente. A la derecha se muestra la figura correcta.

