

4

Aplicaciones de salud para pacientes: metodología de evaluación y validación

Pablo Pérez Huertas
Emilio Monte Boquet
José Luis Poveda Andrés





Pablo Pérez Huertas

Servicio de Farmacia, Hospital Universitario y Politécnico La Fe, Valencia.



Emilio Monte Boquet

Servicio de Farmacia, Hospital Universitario y Politécnico La Fe, Valencia.



José Luis Poveda Andrés

Servicio de Farmacia, Hospital Universitario y Politécnico La Fe, Valencia.

Índice

1. Introducción
2. *App* de salud
3. Evaluación de las *app* de salud
4. Conclusiones
5. Bibliografía

1. Introducción

La salud móvil (*mSalud*, *mHealth*) se define, según la Organización Mundial de la Salud, como la práctica de la medicina y la salud pública con el soporte de dispositivos móviles, como teléfonos, aparatos de monitorización de pacientes, asistentes personales digitales y otros dispositivos inalámbricos¹. Aunque la *mSalud* engloba mucho más que las aplicaciones para dispositivos móviles (*app*), es un hecho que la disponibilidad de estas *app* aumenta de forma imparable en el área de la salud.

El desarrollo de esta nueva tecnología se debe al aumento de la disponibilidad y de la utilización de dispositivos móviles (principalmente *smartphones*) por parte de la población. Así, se estima que en la actualidad ya hay más teléfonos móviles activos que personas en el mundo, y que el 85% de los usuarios accede a Internet desde el móvil². En el caso

de España, el 87% de las líneas móviles son *smartphones* y se estima que 7 de cada 10 españoles que posee un *smartphone* se conecta todos los días a Internet a través de este dispositivo². Además, una reciente encuesta de la Asociación para la Investigación de Medios de Comunicación señala que el 93,9% de las personas que se conectan a Internet en España lo hace a través del teléfono móvil, siendo este el dispositivo más utilizado³.

Otro factor que explica el crecimiento de esta nueva tecnología es el hecho de que cada vez se reconoce más la importancia del paciente activo e informado, dado que de él se esperan mejores resultados en salud y mayores beneficios, tanto individuales como colectivos. Y, sin duda, la *mSalud* puede contribuir a conseguir un mayor empoderamiento en salud de la población.

2. App de salud

En los últimos años se ha producido un desarrollo masivo de *app* relacionadas con la salud. El crecimiento está siendo exponencial; por ejemplo, solo en los Estados Unidos, en el año 2011 estaban disponibles más de 5.800 *app* de salud, en 2012 ya eran más de 13.000 *app* y a partir de 2013 se superaban las 23.000 *app*⁴. Actualmente, algunas publicaciones estiman que se dispone de más de 165.000 *app* de salud en todo el mundo⁵.

La selección de una *app* por parte de los usuarios se realiza mayoritariamente a través de las tiendas de los sistemas operativos (Google Play para Android e iTunes para Apple)⁶. En estas tiendas *online*, las *app* se valoran por dos únicos criterios: el número de descargas y la valoración de los usuarios. Estas plataformas son buscadores inespecíficos que no permiten al usuario aplicar filtros tan importantes como la enfermedad que padecen o la función que quieren que desempeñe la *app*. Este hecho dificulta la búsqueda y hace que el usuario termine descargando *app* que no tienen por qué ser las más adecuadas para sus necesidades. Como segunda opción, los

usuarios seleccionan *app* mediante búsquedas en Internet. Cada vez hay más páginas web dedicadas a ayudar a usuarios a buscar la *app* de salud más adecuada. Este es el caso de myhealthapps.net y de imedicalapps.com, que categorizan las *app* y permiten aplicar filtros específicos a las búsquedas. Así, el usuario puede buscar *app* destinadas a su patología y que contengan una funcionalidad específica, ayudándole a encontrar la que mejor se ajusta a sus necesidades.

Existe una gran variabilidad entre las *app* de salud en lo que se refiere a sus funcionalidades. En un informe emitido por la Comisión europea sobre evaluación de *app* de salud se proponen que se clasifiquen en nueve categorías⁷:

- Ayuda en la toma de decisiones al paciente y herramientas de autogestión.
- Herramientas de soporte de decisiones clínicas para el diagnóstico o el tratamiento.

- Aplicaciones de cambio de estilo de vida y herramientas simples de autocuidado.
- Punto de atención de diagnóstico, seguimiento y tratamiento.
- *App* que controlan dispositivos médicos.
- *App* de comunicación.
- *App* para registro de documentación.
- *App* que facilitan cálculos sencillos.
- Calculadoras médicas genéricas.

La mayor parte de las *app* de salud disponibles se centran en el estilo de vida saludable y tienen como principal función el registro y el seguimiento de ejercicios y el consumo de calorías⁸. En cuanto a las *app* específicas para problemas de salud, las más frecuentes están destinadas al insomnio, la salud mental y el tabaquismo⁸. En un estudio se analizaron las funciones de 116 *app* de salud destinadas a patologías. Las tres funcionalidades más frecuentes fueron información/educación, alertas y recordatorios, y manejo de efectos adversos (Figura 1). Por otro lado, los autores observaron que el 59,5% de las *app* estaban destinadas a pacientes, solo el 12,9% a profesionales sanitarios y el 23,3% para la interacción de ambos colectivos.

El desarrollo masivo de *app* sin una regulación específica ha provocado que se lleguen a difundir *app* de mala calidad, que en el contexto de la salud pueden conllevar problemas serios. Las principales razones de la falta de calidad en las *app* de salud son:

- Escasa evidencia de su efectividad. En un estudio publicado en España realizado

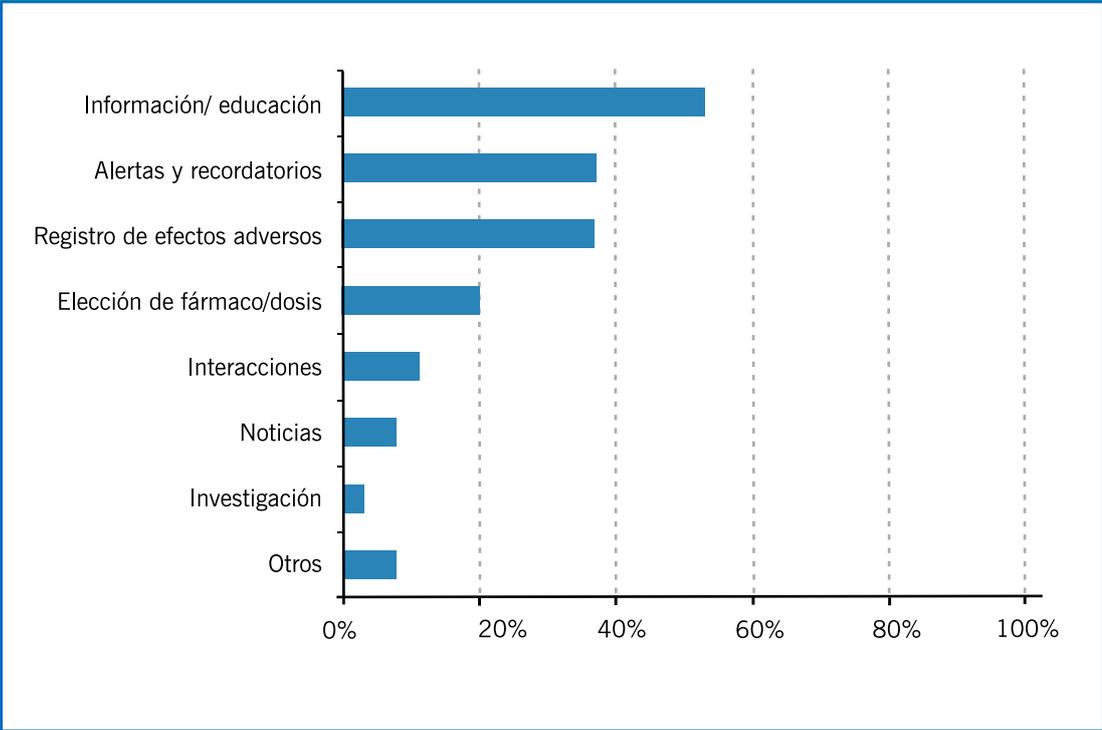
sobre 166 *app* dirigidas a pacientes con cáncer se indica que solo el 50,6% de ellas presentaba una evidencia científica suficiente⁹.

- No funcionan correctamente. Un grupo de investigadores norteamericanos observaron que tres de cada cuatro aplicaciones destinadas a diagnosticar melanomas clasificaban incorrectamente más de un 30% de los casos¹⁰.
- La información que contienen no está validada. En un estudio realizado con *app* sobre hepatitis víricas se evidenció que casi en la mitad de ellas no había ningún profesional sanitario involucrado en su desarrollo, y que la mayor parte de la información contenida no estaba validada¹¹.
- No protegen la confidencialidad y la privacidad de los usuarios. Un estudio realizado sobre 211 *app* destinadas a pacientes diabéticos demostró que el 81% de las aplicaciones no tenían política de privacidad¹².

Se están desarrollando *app* que pueden producir un impacto significativo en la salud de los usuarios, y en estos casos las autoridades reguladoras tienen la responsabilidad de evaluar esta tecnología sanitaria como si de un producto sanitario se tratase y regular su acceso al mercado. En este sentido, la *Food and Drug Administration* de los Estados Unidos publicó en 2015 una guía con directrices para el desarrollo de *app* de salud¹³. Sin embargo, se hace cada vez más necesaria una evaluación objetiva, integral y transparente de las *app* de salud que permita al usuario conocer cuáles cumplen unos mínimos estándares de calidad.



Figura 1. Funcionalidades de las *app* destinadas a problemas de salud⁶.



3. Evaluación de las *app* de salud

Como ya se ha comentado, la evaluación de esta nueva tecnología sanitaria es imprescindible para garantizar la seguridad de los usuarios. Sin embargo, hay muchos interrogantes en relación a esta evaluación: ¿qué criterios son los que miden la calidad de una *app*?, ¿todos los criterios deberían contar de la misma manera?, ¿se deberían evaluar de manera diferente *app* con funciones diferentes?, ¿cómo se debería llevar a cabo esta evaluación para contar con las mayores garantías?, ¿quién debería hacerla?

3.1. Criterios de evaluación

Definir los criterios que deben utilizarse para evaluar las *app* de salud representa el primer paso para desarrollar una metodología de evaluación. Para ello, hay que averiguar a qué características dan más importancia los usuarios que las van a utilizar.

En un estudio se realizó una encuesta a 125 usuarios para conocer qué valoraban más en una *app* de salud, y se observó que las varia-

bles que obtenían mayor puntuación eran la calidad del contenido, la evidencia científica, la seguridad y la privacidad. Por otro lado, las que menor puntuación obtuvieron fueron el tamaño de la *app* y que esté bien identificado el autor¹⁴. En este mismo estudio también se preguntó a los desarrolladores informáticos de las *app* qué criterios consideraban más importantes (Figura 2).

En un estudio realizado sobre pacientes con infección por el virus de la inmunodeficiencia humana¹⁵, los usuarios finales de una *app* identificaron funcionalidades que debería incorporar una *app* ideal para el manejo de sus necesidades de salud: interactividad, posibilidad de compartir información tanto con profesionales sanitarios como con otros pacientes con una situación clínica similar, personalización y que incluya un buscador.

En otra encuesta realizada a 19 universitarios sobre qué consideraban más relevante para una *app* de salud, algunos de los criterios señalados como más importantes fueron la facilidad de uso, la capacidad de mantener el

F02

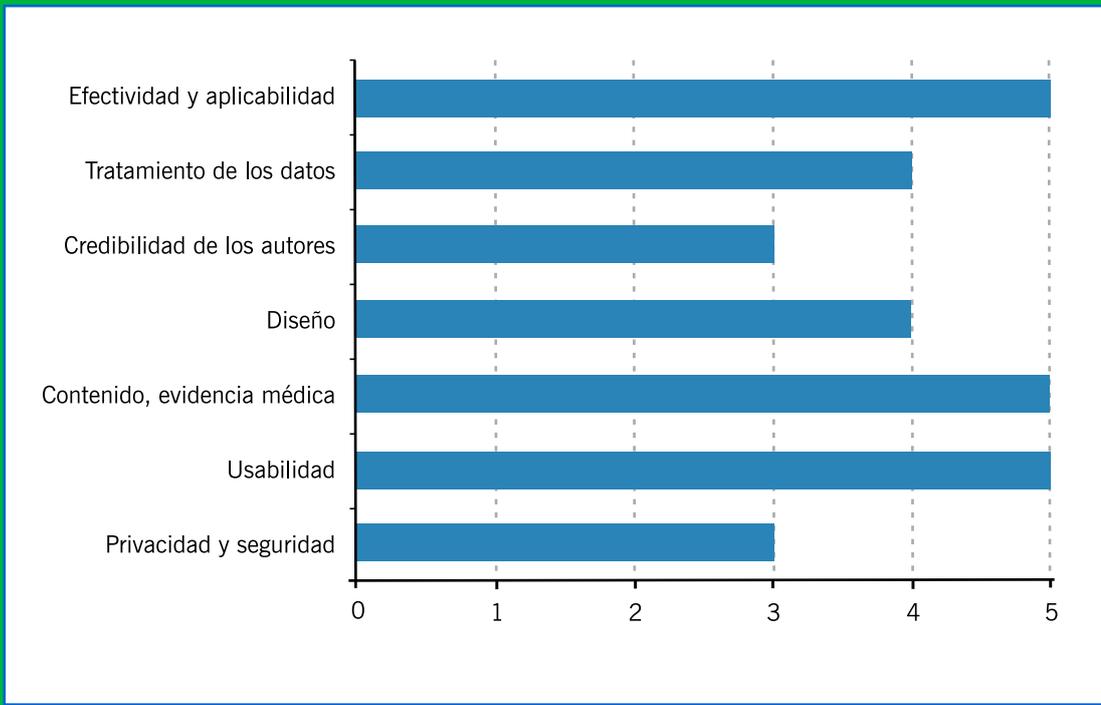


Figura 2. Importancia de los criterios de evaluación para los desarrolladores de *app* (puntuación de 1 a 5)¹⁴.

interés del usuario, el coste, la facilidad para descargarla y configurarla, que esté clara la autoría, la personalización, la validez de la información, y la privacidad y la seguridad¹⁶.

De los estudios comentados, puede resumirse que los usuarios valoran más las *app* de salud sencillas de configurar y de utilizar, con una buena calidad de la información, que sean personalizables y que protejan sus datos. Por otro lado, el aspecto visual y el diseño también deberían reconocerse como factores clave para la utilización de las *app* de salud por parte de los pacientes. Este aspecto es en especial relevante en las personas mayores que, por una parte, tienen más dificultad para utilizar las *app*, y por otra, son un grupo de población diana en mSalud. Un estudio realizado en España demuestra que los pacientes ancianos polimedcados son perfectamente capaces de utilizar una *app* diseñada específicamente para ellos (ALICE), incluso aquellos sin experiencia previa con tecnologías de la información y la comunicación, con buenos resultados en mejora de la adherencia y con un alto grado de satisfacción por parte de los usuarios¹⁷.

3.2. Metodología de evaluación

En la literatura pueden encontrarse distintas metodologías de evaluación de *app* de salud, siendo estas muy variables en cuanto a método de selección y utilización de criterios, sistema de puntuación y aplicación. A continuación se detallan las más destacadas.

3.2.1 AppSaludable

Iniciativa desarrollada a partir de 2012 por la Agencia de Calidad Sanitaria de Andalu-

cía para certificar *app* de salud¹⁸. Incluye recomendaciones para el diseño, el uso y la evaluación de aplicaciones móviles de salud dirigidas a todos los colectivos involucrados: ciudadanía, profesionales sanitarios, proveedores de servicios sanitarios y desarrolladores.

Se definieron 31 recomendaciones seleccionadas por un panel de expertos del que forman parte profesionales sanitarios, desarrolladores de aplicaciones, representantes de pacientes y gestores. Para medir la calidad, recoge recomendaciones sobre el diseño y la pertinencia, la calidad y la seguridad de la información, la prestación de servicios y la confidencialidad y privacidad.

El proceso se basa en la autoevaluación de la aplicación de acuerdo con las recomendaciones de la guía, y la posterior evaluación por parte de un comité de expertos. Una vez se ha concluido el proceso y la *app* cumple con todos los requisitos, obtiene el distintivo AppSaludable y pasa a formar parte de un repositorio en su página web (calidadappsalud.com). Este sello de calidad deberá ser reevaluado periódicamente o cuando se realicen modificaciones significativas.

Actualmente hay 15 *app* que cuentan con el certificado y 28 que están en proceso de certificación.

3.2.2 App Salut

Plataforma desarrollada a partir de 2015 por la oficina mHealth.cat y la fundación TicSalut, pertenecientes al gobierno de Cataluña. Evalúa *app* de salud utilizando cuatro aspectos clave: usabilidad, tecnología, seguridad y contenidos¹⁹. Se trata de un proceso de certificación y evaluación en el cual las *app* que superan unos

requisitos mínimos se difunden a través de un catálogo web (appsalut.gencat.cat).

Un comité de expertos desarrolló 113 criterios de evaluación que se aplican en función del riesgo potencial de la *app* en la población²⁰. Así, las *app* se clasifican en tres niveles de riesgo en función de tres parámetros:

- A. Tipo y cantidad de información sensible de los usuarios que recoge la *app*.
- B. Información o recomendaciones que ofrece al usuario.
- C. Impacto que supone en la población, medido como el número de usuarios potenciales.

Mediante dos matrices de riesgo se asigna un nivel de riesgo tecnológico y otro de contenidos. En función de estos dos niveles, cada uno de los 113 criterios de evaluación se define como no aplicable, deseable, recomendable u obligatorio. Para que una *app* esté acreditada y sea incluida en su catálogo debe cumplir todos los criterios marcados como obligatorios. Una vez superados todos los obligatorios, por cada categoría se calcula la puntuación utilizando la media de los criterios deseables (4 puntos) y los criterios recomendables (6 puntos) que cumpla. La puntuación global es la media global de las cuatro categorías.

La plataforma web acaba de inaugurarse y de momento solo cuentan con dos *app* evaluadas.

3.2.3 Metodología del Consejo Nacional de Información Británico

El Consejo Nacional de Información Británico está desarrollando una nueva plataforma

para evaluar y difundir *app* de salud que sustituirá a la actual *NHS Health Apps Library* (nhs.uk/tools/pages/toolslibrary.aspx). Se trata de una metodología en cuatro etapas:

1. Autoevaluación por parte del desarrollador y posterior auditoría de la información aportada.
2. Evaluación de la comunidad: una red de profesionales de la salud, comisionados y usuarios evalúan las *app* disponibles en la plataforma.
3. Evaluación formal de un reducido número de *app* con mayor calidad identificadas en las etapas 1 y 2.
4. Evaluación del impacto en la población, la eficacia y la rentabilidad.

La idea de los autores es seleccionar un menor número de *app*, pero con mayor calidad para difundir a la población que con el actual sistema.

Los criterios de evaluación que se utilizarán en los pasos 1, 2 y 3 se basan en nueve aspectos clave²¹:

1. Propósito.
2. Planificación y desarrollo.
3. Facilidad de uso.
4. Evaluación inicial y adaptación.
5. Técnicas de cambio de comportamiento.
6. Mantenimiento del comportamiento y prevención de recaídas.
7. Evaluación.
8. Documentación.
9. Protección de datos.

La forma en que se van a llevar a la práctica estos nueve temas no está publicada. En este sentido, McMillan *et al.*²² han propuesto 62 preguntas concretas con dos únicas opciones de respuesta (cumple o no cumple) para aplicar esta metodología.

Al igual que en la metodología AppSaludable, este sistema no otorga una puntuación cuantitativa según su calidad; si la *app* cumple con todos los requisitos descritos se le otorga un sello de calidad y es incluida en su catálogo, y en caso contrario es rechazada.

3.2.4 iSYS Score

Está desarrollada por la fundación iSYS, una organización española destinada a promover proyectos relacionados con la salud y tecnologías 2.0. A partir del año 2015 diseñaron una nueva metodología de evaluación centrada en tres indicadores de calidad: interés popular, confianza y utilidad²³. Solo evalúan *app* destinadas al público general, no a profesionales sanitarios, y la evaluación se basa principalmente en la descripción de la *app* proporcionada por el promotor.

Para definir los criterios de evaluación y su peso específico se llevó a cabo un método Delphi. Se contó con la participación de profesionales sanitarios, pacientes y expertos en salud pública y en nuevas tecnologías. Cada ítem tiene un peso específico según su importancia y se valora de manera dicotómica (lo cumple o no lo cumple). La puntuación por categoría es la suma de los ítems que cumple, y la global es la suma de todos los ítems. La metodología fue validada evaluando 48 *app* de salud escogidas de manera aleatoria.

Anualmente elaboran un *ranking* con las 20 mejores *app* de salud presentadas. Están

disponibles las 20 mejores *app* ordenadas según la puntuación obtenida de los años 2014/2015 y 2015/2016.

3.2.5. Metodología propuesta por Riezebos

Esta metodología ha sido desarrollada por Riezebos, de la Universidad de Ámsterdam, en colaboración con la Universidad de Toronto y la revista *Journal of Medical Internet Research*, y fue publicada en el año 2014¹⁴.

El autor realiza una profunda revisión sistemática sobre la utilización de *app* en salud y busca metodologías validadas de evaluación. Además, lleva a cabo una encuesta para conocer la opinión de los usuarios de las *app* de salud sobre algunos criterios de calidad. A partir de la revisión sistemática y de los resultados de la encuesta, un grupo de expertos extrajo los criterios de evaluación finales. Estos pueden clasificarse en seis categorías:

1. Objetivos y metas.
2. Facilidad de uso.
3. Contenido.
4. Desarrollo.
5. Privacidad.
6. Valoración general.

Esta metodología cuenta con 16 criterios evaluables que se puntúan de 1 a 5 en función del cumplimiento nulo o completo del criterio. Todos los criterios cuentan igual en el resultado final de la evaluación. La puntuación final es la media de las seis categorías.

La evaluación se realiza en tres pasos: el primero incluye una autoevaluación por parte del desarrollador con los criterios que se han definido y una posterior validación por parte de un experto; el segundo paso es una revi-

sión por pares de expertos independientes, y si esta obtiene una resolución positiva, la *app* se difunde a través de su catálogo web (en desarrollo); y el tercer paso tiene lugar una vez que ya es pública la *app* en la plataforma web, cuando los usuarios que la han descargado y la conocen pueden evaluarla utilizando el mismo sistema de puntuación. De esta manera, en la plataforma web las *app* quedan puntuadas por expertos y usuarios, siendo esta última puntuación dinámica. Por otro lado, si los evaluadores lo consideran adecuado, una *app* con buena calidad puede ser elegida para obtener una certificación por parte de los desarrolladores que será reevaluada anualmente.

La limitación de esta metodología es que aún no ha sido validada ni se ha comenzado a utilizar en la plataforma web (reviewtool.hero-kuapp.com).

3.2.6. MARS

Iniciativa desarrollada por miembros del *Institute of Health & Biomedical Innovation* de Brisbane (Australia) y publicada en 2015. Los autores realizaron una revisión sistemática sobre las metodologías de evaluación de *app* o de páginas web de salud para buscar los criterios más utilizados y su aplicabilidad²⁴. Reunieron 349 criterios diferentes, de los que un grupo de expertos extrajo 23 criterios de evaluación. Se clasifican en cinco categorías:

1. Compromiso.
2. Funcionalidad.
3. Estética.
4. Calidad de la información.
5. Calidad subjetiva.

Los 23 criterios de evaluación se puntúan de 1 a 5 según su grado de cumplimiento. Cada

ítem cuenta con cinco opciones de respuesta diferentes según el criterio que se está evaluando, de manera que se facilita la decisión del evaluador. Además, se ha añadido la posibilidad de dejar un ítem sin contestar en caso de no ser aplicable a la *app* evaluada. La puntuación de cada ítem tiene el mismo peso en la puntuación final. La puntuación global es la media de todos los criterios excepto el último (calidad subjetiva), que se informa de manera separada. En caso de que hubiera algún ítem no aplicable, se extraería del cálculo. La metodología fue validada evaluando 50 *app* de salud mental.

Recientemente la metodología MARS ha sido rediseñada para su aplicación por parte de los usuarios. Para ello, se ha simplificado y se ha adaptado el lenguaje. La nueva metodología uMARS también fue validada por 164 usuarios²⁵.

MARS y uMARS se han utilizado para evaluar cientos de *app* de salud. También se han empleado en estudios independientes de evaluación^{26,27}. Sin embargo, no existe una plataforma web donde se difundan las *app* evaluadas con las puntuaciones obtenidas.

3.3. Aplicación

Del análisis de las diferentes metodologías comentadas se ve claramente que todas parten de un mismo objetivo principal, pero lo llevan a la práctica de manera muy diferente. En la Tabla 1 se resumen las principales características de cada metodología.

El origen de los criterios de evaluación es, sin duda, uno de los aspectos más importantes al estudiar la metodología de evaluación. Como puede observarse en la Tabla 1, las tres me-

T01

Tabla 1.

Características principales de las metodologías de evaluación de *app* de salud.

| | Diseño de criterios | Criterios ponderados | Tipo de puntuación | Validada | Resultado |
|--------------|--|--------------------------------|--------------------|----------|---------------------------------------|
| AppSaludable | Consenso expertos | No | Sí/no | No | Autoevaluación y certificación |
| App Salut | Consenso expertos | En función del nivel de riesgo | Sí/no | No | Certificación y lista |
| Británica | Consenso expertos | No | Sí/no | No | Autoevaluación y certificación |
| iSYS Score | Método Delphi | Sí | Sí/no | Sí | Lista |
| Riezebos | Revisión sistemática y encuesta a usuarios | No | 1-5 | No | Autoevaluación, lista y certificación |
| MARS | Revisión sistemática | No | 1-5 | Sí | Lista |

metodologías públicas están diseñadas por consenso de expertos. En el caso de iSYS Score se recurre a un método Delphi que se basa en un consenso de expertos, pero con una metodología sistemática que permite extraer conclusiones con mayor validez. La metodología cuyos criterios de evaluación han sido más estudiados ha sido la de Riezebos. Para ello, el autor realiza un profundo estudio de las metodologías de evaluación de nuevas tecnologías sanitarias y de las encuestas realizadas a usuarios. Además, también lleva a cabo un estudio para conocer la opinión de los usuarios y de los desarrolladores.

En cuanto al contenido de los criterios de evaluación, la Comisión Europea publicó un informe en el que resume y combina las diferentes metodologías de evaluación de *app* de salud publicadas. Como puede verse en la Tabla 2, las diferentes metodologías coinciden en la mayoría de los aspectos clave a evaluar, aunque se diferencian en cómo se traducen en criterios concretos.

En el mismo informe de la Comisión Europea⁷ también se recogen las cuatro formas de evaluar *app* de salud:

- *Lista*: utilizar una lista sencilla de preguntas que puede disponer de una puntuación variable o fija. Es un sistema sencillo de aplicar y cuantitativo que facilita la comparación de *app*.
- *Tanteo*: en función del nivel de riesgo que presenta una *app*, la puntuación de cada criterio individual de una lista de preguntas presentará una puntuación diferente. Es más complejo que la lista, pero ofrece resultados cuantitativos que se adecuan más a la realidad.

- *Autoevaluación*: se realizan preguntas directas a los desarrolladores y estos exponen sus argumentos. Posteriormente la información es evaluada de manera independiente. Se utiliza como una preevaluación en los procesos de certificación.
- *Certificación*: si la *app* cumple una serie de requisitos estrictos se le otorga un sello de calidad, que debe ser reevaluado al transcurrir un determinado tiempo o al realizar modificaciones significativas. Es un sistema cualitativo utilizado por instituciones públicas como sello de calidad.

En función del objetivo de la evaluación y del destino de esta, se deberá aplicar un método u otro. Como hemos visto, las instituciones públicas, en sus procesos de certificación, utilizan la autoevaluación como primera evaluación para luego aplicar unos criterios estrictos de certificación. App Salut es la única que aplica el tanteo para evaluar en función del riesgo de la *app*. Este sistema es muy interesante ya que, por ejemplo, el resto de las metodologías evalúan de la misma manera una *app* destinada a buscar la farmacia más cercana que otras que ayudan a calcular la dosis de insulina a pacientes diabéticos, cuando el impacto de la calidad sobre los usuarios es muy diferente. Además, App Salut incorpora un sistema de puntuación para poder comparar *app* que hayan sido certificadas. Esto resulta muy útil a los usuarios porque puede orientarles para seleccionar *app* de mayor calidad.

Otro aspecto muy importante es el sistema de puntuación. En este sentido, hay dos preguntas a discutir: ¿todos los criterios deberían tener el mismo peso en la puntuación final? y

T02

Tabla 2.

Criterios de evaluación utilizados en las metodologías estudiadas. Adaptada del Informe de la Comisión Europea⁷.

| | AppSaludable | App Salut | iSYS Score | Británica | Riezebos | MARS |
|-------------------------------------|--------------|-----------|------------|-----------|----------|------|
| Utilidad y deseabilidad | | | | | | |
| Pertinencia | + | | | | | |
| Accesibilidad | + | + | | + | | |
| Diseño | + | + | | | + | + |
| Usabilidad | + | + | + | + | + | + |
| Gamificación | | | + | | | |
| Adecuación a la audiencia | + | | | | + | + |
| Entretenimiento | | | | | | + |
| Interacción | | + | + | | | + |
| Precisión en la descripción | | + | | | | + |
| Credibilidad | | | | | | |
| Credibilidad de los autores | + | + | + | | + | + |
| Actualización | + | | | | + | |
| Fuentes de información | + | + | + | + | + | + |
| Interoperable y transparente | | | | | | |
| Interoperabilidad | | | | + | | |
| Información abierta | | | | + | | |
| Transparencia | + | | + | + | | |
| Fiabilidad técnica | | | | | | |
| Fiabilidad/replicabilidad | | + | | | | + |
| Soporte técnico | + | + | | + | | + |
| Ancho de banda | + | | | | | |
| Seguridad | | | | | | |
| Gestión de riesgos | + | + | | + | | |
| Efectividad | | | | | | |
| | | | | + | + | + |
| Privacidad y seguridad | | | | | | |
| Protección de datos | + | + | + | + | + | |
| Seguridad del software | + | + | | + | + | |
| Otros | | | | | | |
| Comercio electrónico | + | | | | | |
| Publicidad | + | | | | | |
| Manejo de datos | | | | | + | |
| Puntuación usuarios | | | + | | | |
| Compatibilidad plataformas | | | + | | | |
| Adaptabilidad tecnológica | | + | | | | |

¿los criterios deberían puntuarse de manera cualitativa o cuantitativa? En el caso del método MARS, por ejemplo, todos los criterios cuentan igual para la puntuación final. Nos encontramos con que es igual de importante que una *app* sea agradable a la vista y que la información que ofrezca cuente con un nivel de evidencia científica adecuado. Resultan más reales las escalas en las que se pondera cada criterio en función de su relevancia. El método de App Salut sigue siendo el más completo en este sentido. La iSYS Score también puntúa cada variable según la relevancia en la evaluación.

Por otro lado, los métodos que valoran los criterios de manera cualitativa (lo cumple o no lo cumple) pueden resultar menos aplicables en la práctica. Por ejemplo, la primera recomendación de App Salut es “Los elementos (texto, imágenes, iconos, botones, etc.) son identificables y fáciles de utilizar”, y la valoración de este criterio es dicotómico aunque

la interpretación abarca una amplia variedad de posibilidades. Puede tener una tipografía muy adecuada, pero los iconos ser un poco pequeños y los botones no estar fácilmente localizables; y estos casos intermedios pueden generar variabilidad en los resultados y mayor dificultad para los evaluadores. Las metodologías que ofrecen rangos de puntuación tienen una mejor interpretación. En el caso de MARS, los criterios son preguntas concretas para las que se ofrecen cinco opciones de respuesta en función del criterio preguntado, lo que facilita la evaluación y reduce la variabilidad interoperador.

El último aspecto a tener en cuenta en las metodologías es la validación. Como en cualquier otra metodología, la validación debe ser un requisito indispensable para asegurar la validez de los resultados. En la Tabla 1 se recoge que solo hay publicados estudios de validación de iSYS Score y de MARS.

4. Conclusiones

El gran número de *app* de salud disponibles en el mercado y su impacto sobre la población hacen necesario que se realice una evaluación transparente, objetiva e integral de ellas por parte de los expertos en mSalud y de los usuarios.

Las seis metodologías estudiadas coinciden en los principales aspectos que definen la calidad de una *app* de salud: navegabilidad, utilidad, privacidad y confidencialidad y criterios tecnológicos. Se observa una mayor variabilidad en cómo traducen las categorías en criterios de evaluación concretos y en el desarrollo de sistemas de evaluación y puntuación.

La metodología más adecuada para utilizar en catálogos web donde los pacientes puedan buscar y comparar *app* de salud es la MARS. Representa un sistema de evaluación

sencillo, cuantitativo y validado, que permite una evaluación rápida y con poca variabilidad. Además, gracias a su versión para usuarios (uMARS) también permitirá la evaluación en paralelo por los usuarios.

Por otro lado, en el ámbito de la certificación realizada por instituciones públicas, la metodología App Salut es la más completa. Representa la única que evalúa en función del potencial impacto sanitario mediante una matriz de riesgo. Además de la certificación, añade una evaluación cuantitativa según la calidad adicional de la *app* que permite al usuario orientar su búsqueda hacia aquellas de mayor calidad.

Por último, se requieren más estudios para comparar las diferentes metodologías, así como para conocer el impacto de la evaluación de las *app* de salud sobre los usuarios.

5. Bibliografía

1. WHO. Global Observatory for eHealth Series - Volume 3: mHealth: new horizons for health through mobile technologies. 2011.
2. Ditrendia. Informe Ditrendia: mobile en España y en el mundo 2015. Disponible en: <http://www.ditrendia.es/wp-content/uploads/2014/07/Ditrendia-Informe-Mobile-en-Espa%C3%B1a-y-en-el-Mundo.pdf>
3. AIMC. 18º Informe Navegantes En La Red. Encuesta a Usuarios de Internet.; 2016.
4. Kamerow D. Regulating medical apps: which ones and how much? *BMJ*. 2013; 347:f6009.
5. IMS Institute for Healthcare Informatics. Patient adoption of mHealth. Use, evidence and remaining barriers to mainstream acceptance. 2015.
6. van Kerkhof LWM, van der Laar CWE, de Jong C, Weda M, Hegger I. Characterization of apps and other e-tools for medication use: insights into possible benefits and risks. *JMIR mHealth uHealth*. 2016;4(2):e34.
7. European Commission. Public consultation on the Green Paper on mobile Health. 2016.
8. Sama PR, Eapen ZJ, Weinfurt KP, Shah BR, Schulman KA. An evaluation of mobile health application tools. *JMIR mHealth uHealth*. 2014;2(2):e19.
9. Collado-Borrell R, Escudero-Vilaplana V, Ribed-Sánchez A, Ibáñez-García S, Herranz-Alonso A, Sanjurjo-Sáez M. Smartphone applications for cancer patients; what we know about them? *Farm Hosp*. 2016;40(1):25-35.
10. Wolf JA, Moreau JF, Akilov O, *et al*. Diagnostic inaccuracy of smartphone applications for melanoma detection. *JAMA Dermatology*. 2013;149(4):422.
11. Cantudo-Cuenca MR, Robustillo-Cortés MA, Cantudo-Cuenca MD, Morillo-Verdugo R. A better regulation is required in viral hepatitis smartphone applications. *Farm Hosp*. 2014;38(2):112-7.
12. Blenner SR, Köllmer M, Rouse AJ, *et al*. Privacy policies of Android diabetes apps and sharing of health information. *JAMA*. 2016;315(10):1051.
13. U.S. Department of Health and Human Services. Food and Drug Administration. Mobile medical applications. Guidance for Industry and Food and Drug Administration Staff. 2015.
14. Riezebos RJ. Peer-reviewing of mHealth applications. 2014. Disponible en: <http://dare.uva.nl/cgi/arno/show.cgi?fid=573074>
15. Schnall R, Mosley JP, Iribarren SJ, Bakken S, Carballo-Diéguez A, Brown Iii W. Comparison of a user-centered design, self-management app to existing mHealth apps for persons living with HIV. *JMIR mHealth uHealth*. 2015;3(3):e91.
16. Dennison L, Morrison L, Conway G, Yardley L. Opportunities and challenges for smartphone applications in supporting health behavior change: qualitative study. *J Med Internet Res*. 2013;15(4):e86.
17. Mira JJ, Navarro I, Botella F, *et al*. A Spanish pillbox app for elderly patients taking multiple medications: randomized controlled trial. *J Med Internet Res*. 2014;16(4):e99.

18. Agencia de Calidad Sanitaria de Andalucía. AppSaludable. (Consultado el 13/7/2016). Disponible en: <http://www.calidadappsalud.com/>
19. App Salut. Guía de criterios del proceso de acreditación. 2016. Disponible en: <http://www.ticsalut.cat/media/upload//arxius/portal-appsalut/Gu%C3%ADa%20criterios%20proceso%20acreditaci%C3%B3n.pdf>
20. App Salut. Clasificación por niveles y coste económico. 2016
21. National Institute for Health and Care Excellence. Behaviour change: general approaches. (Consultado el 13/7/2016.) Disponible en: <https://www.nice.org.uk/guidance/ph6>
22. McMillan B, Hickey E, Patel MG, Mitchell C. Quality assessment of a sample of mobile app-based health behavior change interventions using a tool based on the National Institute of Health and Care Excellence behavior change guidance. *Patient Educ Couns*. 2015;99(3):429-35.
23. Grau I, Kostov B, Gallego JA, Grajales III F, Fernández-Luque L, Sisó-Almirall A. Método de valoración de aplicaciones móviles de salud en español: el índice iSYS Score. *Semer - Med Fam*. 2016.
24. Stoyanov SR, Hides L, Kavanagh DJ, Zelenko O, Tjondronegoro D, Mani M. Mobile App Rating Scale: a new tool for assessing the quality of health mobile apps. *JMIR mHealth uHealth*. 2015;3(1):e27.
25. Stoyanov SR, Hides L, Kavanagh DJ, Wilson H. Development and validation of the user version of the Mobile Application Rating Scale (uMARS). *JMIR mHealth uHealth*. 2016;4(2):e72.
26. Masterson Creber RM, Maurer MS, Reading M, Hiraldo G, Hickey KT, Iribarren S. Review and analysis of existing mobile phone apps to support heart failure symptom monitoring and self-care management using the Mobile Application Rating Scale (MARS). *JMIR mHealth uHealth*. 2016;4(2):e74.
27. Domnich A, Arata L, Amicizia D, *et al*. Development and validation of the Italian version of the Mobile Application Rating Scale and its generalisability to apps targeting primary prevention. *BMC Med Inform Decis Mak*. 2016;16(1):83.