European Commission

First draft of guidelines

EU guidelines on assessment of the reliability of mobile health applications

First draft of guidelines

A study prepared for the European Commission
DG Communications Networks, Content & Technology by:

CONSARD LTD

Digital Agenda for Europe

This study was carried out for the European Commission by

Consard Limited

Andrew Ruck

Susie Wagner Bondorf

Charles Lowe

Internal identification

Contract number: 30-CE-0763782/00-20

The mHealth app market in Europe is facing challenges. In order to tackle these, the European Commission launched a public consultation entitled "Green Paper on mobile health" in April 2014 asking stakeholders their views. The results were published in January 2015. Privacy emerged as the single most important issue concerning users, resulting in an initiative to develop a code of conduct for the privacy of mHealth apps, now close to completion.

Safety and transparency of information were also identified as key issues along with data quality when linking mHealth apps to Electronic Health Records (EHR) for the effective uptake in clinical practice. A number of stakeholder meetings were organised during 2015, and the outcome was a common understanding that there are health and safety risks related to mHealth apps which need to be handled with regards to:

1. Clinical evidence;

2. Claims on the purpose and functions of mHealth apps;

3. Test and validation of the performance.

The European Commission has appointed a Working Group, facilitated by a Development Leader, Consard Limited, to progress the development of common EU guidelines using an agreed assessment methodology, covering the above mentioned aspects.

This current version is the first draft. Structured consultation is now planned with a range of stakeholders: for more information about how to get involved, please contact CNECT-MHEALTH-EXPERTGROUP@ec.europa.eu .

Feedback from this consultation will lead to changes & refinement of the contents and of the approach to the Guidelines. At least three more drafts are anticipated, with further consultation and a feedback taking place on each iteration of the guidelines. A final version of the guidelines is anticipated by the end of 2016.

INTRODUCTION

The purpose of the mHealth assessment guidelines is to establish a framework of safety, quality, reliability and effectiveness criteria to improve the use, development, recommendation and evaluation of mHealth apps. This with the clear goal to facilitate prevention and an overall healthcare advancement through a controlled use of mobile technology.

It is proposed that the target groups to whom the guidelines are aimed will be:

1. Citizens
2. App developers
3. Healthcare professionals
4. Private and public healthcare providers, public authorities and health insurance providers

To explain the main expected benefits a "do nothing scenario" for each group is listed here:

1. Citizens
    - Lack of trust might cause low end user utilisation of apps.
2. App developers
    - Europe might be a less favoured place for mHealth business because of poor market conditions.
3. Healthcare professionals
    - No joined up service provision because available apps may not be suited to their immediate environment or take account of specific clinical needs.
4. Private and public healthcare providers, public authorities and health insurance providers
    - They will have to devote effort and resources to developing their own guidelines, causing risk of duplication and conflicting guidelines country by country.

To summarise in a sentence, the aim of the guidelines is “better use of better apps for better healthcare”. It seems probable that different but internally consistent versions of the guidelines could be produced, suitable for the above audiences (target groups). Similarly, a range of means of publication is probable: for instance brochures, web-pages and more comprehensive documents.

The developed guidelines are voluntary. In parallel, explicit linkages to existing EU or MS legislation or regulation should be developed to ensure regulatory compliance.

Expected audiences for the guidelines

Those expected to benefit most from the guidelines are (in priority order):

- Citizens / patients / consumers / carers /end-users
- App developers

- Business / Health Business owners / Industry / Manufacturers
- Clinicians / Healthcare professionals
- Healthcare providers / organisations
- Assessment bodies / Regulators
- Authorities / Public administration
- Funders of apps and healthcare, e.g. health insurance

The types of organisations expected to be implanting the guidelines (in priority order) are:

- App developers
- Authorities / Public administration / EC / MS Governments
- Business / Health Business owners / Industry / Manufacturers/ Vendors
- Healthcare providers / organisations
- Citizens / patients / consumers / carers /end-users / consumer counsellors or advocacy
- Assessment bodies / Regulators
- Professional associations
- Policy makers
- Funders of apps and healthcare, e.g. health insurance
- Clinicians / Healthcare professionals

The way these organisations would apply these guidelines in practice would be by means of:

- Dissemination & Promotion
- Development / specification of tools
- Evaluation of apps against Quality criteria
- Legislation / regulation
- Integrate into assessment methodologies (Quality MS) and audits
- Certification / labelling
- Tailored recommendations to eg stakeholders organisations, professional bodies and patient associations
- Linkage of app data to electronic health records
- Support for management of patients / case loads

The guidelines could be tailored to different potential target groups (in priority order) by means of:

- Audience specific documentation and messages / show audience specific benefits, use cases
- Divide / separately address different areas of scope or categories;
- No need for tailoring - one size needs to fit all

- Consultation / concertation
- Long and short version

SCOPE

The three main groups of apps that the term "mHealth apps" is considered to apply to are:

1. CE marked medical devices (not to be covered by these guidelines)
2. Other apps used in a medical setting
3. Health & wellbeing apps

Breaking down the second and third categories above, the specific types of apps covered by the guidelines, in priority order, are:

a. Patient/carer decision aids & self-management tools
b. Clinical decision support tools for diagnosis/treatment recommendation
c. Behaviour change apps - simple self-management tools
d. Point-of-care diagnosis, monitoring or treatment aids
e. Access & editing of EHRs
f. Apps that control medical devices
g. Communication apps - eg teleconsultation
h. Apps providing documentation functionality &/or display a simple measurement
i. Registries & vital events tracking - public health surveillance
j. Simple calculators of on-personal information (eg BMI)
k. Generic medical calculators

The following types of apps are considered to be out of scope because the risk level is too low to merit the assessment detail proposed:

- Apps with view-only functionality
- Apps delivering administrative functions - eg appointment scheduling
- Social forum - networking opportunities

A fourth app type - Apps that control medical devices - can also be excluded as, being accessories to medical devices, these would be covered by medical device legislation too.

mHealth Guidelines vis a vis EU medical devices legislation

The guidelines should make clear that health apps falling under the medical device definition are covered under the medical devices legislation. Since reliability and validity of these apps is addressed through the medical device CE certification process, these guidelines will not specifically address requirements for apps covered under the medical devices legislation. Regulatory compliance should be one of the aspects to be considered in the assessment.

There is a need to deal with the "grey zone" as the delineation with the medical devices is not always clear. The criteria for those apps that are on the borderline and could fall under the medical device definition could be aligned with the medical devices requirements as far as possible.

Therefore, for safety purposes, where "health apps" may create a hazardous situation, they should be treated -in terms of development scrutiny, documentation, verification, validation, etc., similar to medical devices.

Points of agreement include:

- What the FDA describes as the 'grey zone', of apps that technically meet the FDA's definition of medical devices though that they aren't proposing to regulate – in the EU this is interpreted as those apps that are close to being classified as medical devices, so these would be included in the scope of these guidelines.

- 'Off-label' applications where apps with other intended purposes were being used for medical purposes for which they were not originally intended – it is agreed that only the apps' intended uses should be considered in the guidelines to be produced.

- It is important to include apps aimed at prevention.

GUIDELINES

Structure

All app evaluation systems that have or are being developed in the EU comprise a set of separate activities (although these may not be explicitly recognised with separate activities). These are:

- Initial validation – that the app exists, is appropriate for the evaluation, is downloadable etc.
- Risk assessment – which in turn determines the appropriate level of scrutiny
- Scrutiny – of both the technological and the medical aspects

This is the structure proposed for this initial draft.

Note that some systems also seek separately to quantify efficacy – not proposed for this initial set of guidelines.

Initial information gathering & validation: questions for the developer/supplier

1. App name
2. Supplier
3. Developer (if different from (2))
4. Is the app CE certified as a medical device? (if 'yes' terminate assessment)
5. Is app primarily health or social care?
6. Which of the following categories does the app fit into (indicate all that apply):
    a. Patient/carer decision aids & self-management tools

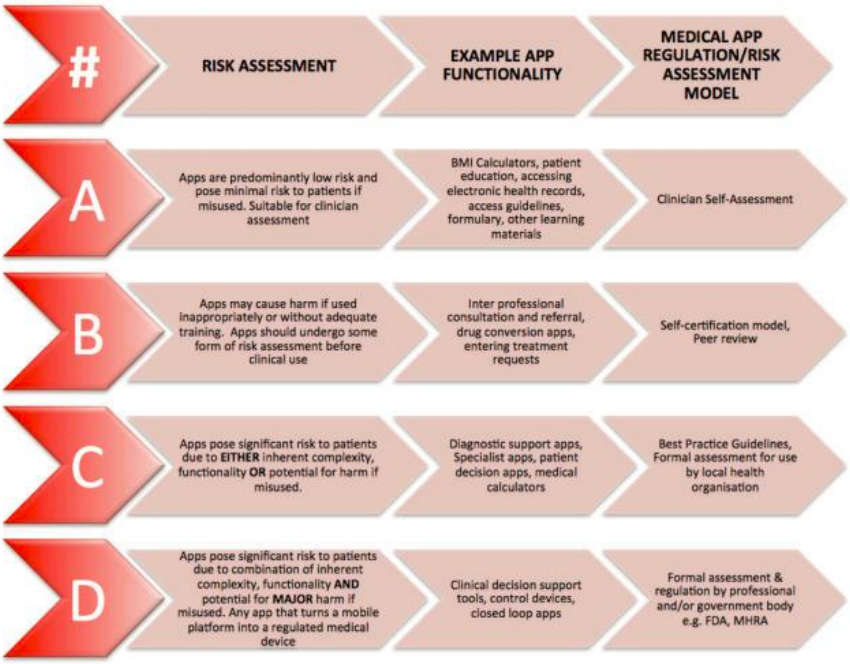
   b. Clinical decision support tools for diagnosis/treatment recommendation
   c. Behaviour change apps – simple self-management tools
   d. Point-of-care diagnosis, monitoring or treatment aids
   e. Access & editing of EHRs
   f. Apps that control medical devices
   g. Communication apps – e.g. teleconsultation
   h. Apps providing documentation functionality &/or display a simple measurement
   i. Registries & vital events tracking – public health surveillance
   j. Simple calculators of on-personal information (eg BMI)
   k. Generic medical calculators

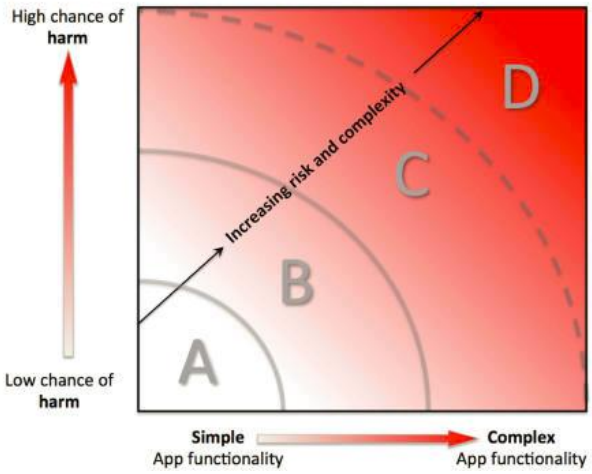
   (if it is a medical app and it does not fit any of the above, terminate assessment)
7. What is the intended use?
8. Please give brief functional description:
9. Please provide academic references for the principles underlying the functioning of the app:
10. Who are the principal beneficiary/ies? (indicate all that apply)
    a. Patient
    b. Carer
    c. Professional user
    d. Healthcare provider
11. How many users have tested the app? (if >one type of user, please give breakdown)
12. Is the app covered by the EU voluntary code on mHealth app privacy?
13. What platforms is the app available on?
14. Please give a brief technical description:
15. What steps have been taken to validate the operation of the app on each platform?

Initial test

- Install/uninstall app on each available platform
- For each platform:
  - Is it easy to understand?
  - Are the screens easy to navigate?
  - Check basic operation: does it work as stated?

Risk assessment





There is a separate discussion underway on risk assessment, yet to be finalised. When completed the intention is to use the clinical risk and the technological risk matrixes to drive the degree to which the answers to the following questions are considered relevant/essential (see Section E below for an illustration). Alternatively, a variation on the approach proposed by Lewis & Wyatt, shown here, might be considered more appropriate.

Scrutiny

Note that whilst many of these questions are equally at home in assessments of non-medical apps, they are nevertheless vitally important in the current context as, for example an app that is not desirable will not continue to be used long by patients.

Is the app usable & accessible?

NB – to be checked separately on every platform offered

1. Is the registration form easy to complete quickly? *(Do we want to limit the number of fields?)*
2. Is the registration form format simple and open (unrestricted characters, numbers, uppercase, etc.)?
3. Do the registration fields incorporate support mechanisms to facilitate the process (pre-determined schedules, scroll down menu, descriptions, etc.)?
4. Are all the separate elements of the app (text, images, icons, buttons, etc.) identifiable and easy to use?
5. Are the colours of the elements appropriately contrasted with the background, (eg avoid similar red/green/brown colour intensities)?
6. Is the text easily readable (size, colour, font) & understandable?
7. Do controls, objects, icons and images have text tags to indicate their function or meaning?
8. Are there visual or vibration alternatives to warning sounds?
9. Does it accept & show all appropriate international characters correctly?
10. Does accessing the service (sending an email confirmation, validation of data access, etc.) happen quickly?
11. Does it fit within the standard interface of a typical mobile device?
12. Are the steps to follow clear; do they make sense?
13. Is there a navigation menu that provides direct access to all functionalities of the app?
14. Is navigation within the app easy & is it clear where in the app the user is?
15. Is it easy to go back to Home directly, and to return to the previous screen?
16. Can the user access any function in the app within three steps? *(Or do we want to make that more?)*
17. When inputting information, is it clear which fields remain to be completed, or are incorrect?
18. Is there access to self-help, video tutorials, guides and FAQ sections to help users?
19. Are there helplines (email, phone, contact form) readily available to resolve questions, problems or incidents?
20. Do the required direct inputs (GPS, sensors, peripherals etc.) work properly?
21. Do the separate functions incorporated in the app load quickly, within a reasonable time?

22. Is the function of each element of the app obvious (clickable, static, drop down, selector, video, etc.)?
23. Are these elements in (22) appropriately positioned & sized to be intuitive, readable and effective to use?
24. Are the visual icons understandable; do they clearly reflect their associated functionality?
25. Is the keyboard used suitable for each type of entry?
26. Where there is a short timeout for screens, is the reading time sufficient?
27. Where the same app is available on different platforms, is the usability experience similar?

Is the app desirable to use?

28. Is the visual identity of the logo in harmony with the visual pattern of the application?
29. Is colour coding uniform and aesthetically pleasing?
30. Are all the graphic elements (pictures, icons, buttons, etc.) used in the same way in all views, consistently?
31. Do the visual icons make the app attractive?
32. Are there any obvious usability problems? (e.g. a button on a device too small to be pressed)
33. Is audiovisual and textual content combined in a balanced & appealing way?
34. Is the color scheme is balanced, not using any particular colour excessively?
35. Is the application properly localised for each country in which it is to be used; is the language/choice of languages appropriate, the currency correct etc.?
36. Is each language used correctly, with no spelling or grammatical errors?
37. Does it follow the interface user guidelines of the operating system?

Is the app credible?

38. Has it been validated by an appropriate group of specialised professionals, health organisation or scientific society?
39. Does it Indicate the sources of information of the contents listed?
40. Does it provide references to the scientific evidence used to ensure content quality?
41. Is there appropriate information provided about the authors of the app content to generate credibility and provide quality assurance?
42. Does it indicate how often the app's content is reviewed/updated?
43. Does it indicate the last review date?
44. Does it notify changes/modifications made at the last update?

Is the app transparent?

45. Does it use simple and understandable language, with clear and short messages, adapted to the target user profile in terms of style and comprehension level?
46. Does it clearly identify who holds any personal data?

47. Does it clearly identify any organisations other than the supplier who have collaborated on the development of the app?
48. Is there concise information on the procedures used to select the app's contents?
49. Does it clearly identify who is/are responsible for the contents of the app?
50. Is there sufficient information on the funding sources, promotion and sponsorship of the app?

Is the app reliable?

51. Does the application logo relate to the purpose of the app?
52. The language change works and is adjusted properly to the interface and contents
53. It is able to properly handle problems with the device and errors of precision, hardware, or from an inadequate use.
54. Does it Inform the user if it requires a long boot up time (default < than 5 seconds)?
55. Dies it notify the user where there is a lengthy operation?
56. Does it allow the user to cancel lengthy operations?
57. Does it notify the user in the case of an external interruption (e.g. loss of network connectivity, database problem)?
58. Does it notify the user in the case of a low bandwidth network?
59. Does it indicate which mobile devices it will work with satisfactorily (according to the operating system, screen resolution, etc.)?
60. Does the screen refresh work properly on the device, including orientation changes, pop-up menus, pop-ups, etc.?
61. Is the information architecture of the application symmetrical, harmonious and proportionate?
62. If the user accepts an incoming call while the application is running, is it possible to return to the same point at the end of the call?
63. Does it behave appropriately in real conditions outside the laboratory?

Is the app technically stable?

64. Does it reject & warn of clearly erroneous data inputs (formats, ranges, etc.)?
65. Is it resilient to abrupt failure during use (locks, etc.)?
66. Is it resilient to changes in other apps, and to external interrupts (incoming call, receiving a message, etc.)?
67. Does it always only consume acceptable levels of resource: battery, CPU, memory, etc.?
68. Does it avoid ever using excessive network resources?
69. Does the app install and uninstall properly?
70. Does it performance remain at the same level in spite of prolonged usage?
71. When the application runs in the background does it do so without affecting other applications or system functions, unless it is specifically designed to do so?

72. Are the database resources appropriately shared between the application and the operating system?
73. Is the application speed acceptable for the purpose required without modifying the user experience or becoming uncontrollable?
74. Does it fail under high load or demand service?
75. Is it able to continue working correctly if repeatedly suspended and resumed?
76. Is it able to continue working correctly if network availability is intermittent?
77. Can it operate (albeit at reduced functionality) in airplane mode, or otherwise with loss of network connectivity?
78. If it requires regular interaction with the user, does it resume successfully from a suspended state at the agreed time/date of each diaried interaction?

Is the app safe?

79. Does it advise that the app is not intended to replace relevant professional services?
80. Does it warn of the possible risks if the app is misused?
81. Does it warn of possible adverse risks caused by the use of the app?
82. Does it provide appropriate guidance if it handles information/data about minors?
83. Does it provide appropriate guidance if it handles information/data about a dependent person who is not the user?
84. Is the supplier's cookie policy stated, and clear?
85. Are there persistent relevant warnings, until the user provides important information or accepts output information?

Is the app effective?

86. Do the functions incorporated provide value to users, in terms of saving time/money, improving information or better health/care?
87. Is it clear who the targeted users are for the app?
88. Is it clear what the intended benefits are to those users?
89. Are the contents and functions offered of potential interest for the user profile to which the app is addressed?
90. Is it clear how those users will need to change the care pathway they participate in (if professional), or lifestyle, in order best to benefit from the app?
91. Is this change (in (4)) realistically achievable?
92. Does it evidence real benefit to users?
93. Has that benefit been evidenced acceptably?

Is the app private & secure?

Note for this section, preferably, we could merely specify compliance with the EU Privacy Code of Conduct for mHealth apps.

94. Is it clear if user registration is necessary for full operation?

95. Is it clear to the user what user data is collected by the app?
96. Is it clear to the user why the data is being collected, by whom and for what purpose?
97. Is it clear to the user whether the data collector will do anything else with the user's personally identifiable data?
98. Is it clear to the user whether the data collector will do anything else with the user's data appropriately anonymised?
99. If third parties have access to data, is this in an acceptable manner, with user approval only?
100. Does it describe the app's maintenance policy for storage & deletion of data provided by the user?
101. Are user data authentication processes acceptable?
102. Does it describe the rights of access, rectification, cancellation or removal of personal data?
103. Can it be confirmed that passwords are not stored directly on the device?
104. Does it manage access to the user's personal information appropriately, with user approval?
105. Are the permissions requested to access the different services of the device clearly described?
106. Are the communication channels used appropriately encrypted?
107. Are the mechanisms of authorisation and authentication adequate?
108. Is the app source code inaccessible & unalterable by the user?
109. Does the app comply with the GDPR principle of *data minimisation*?
110. Does the app comply with the GDPR principle of *data protection by default*?
111. Does the app comply with the GDPR principle of *data protection by design*?
112. If the app is able to write personal information to a patient's electronic health record does it comply fully with the EHR provider's interoperability and security requirements?

METHODOLOGY

As yet no decision has been taken by the Working group on how to apply the assessment criteria.

The options include:

Scoring

This involves calculating a risk-related score for each app, with a cut-off below which the app is rejected, plus some questions for any of which the answer 'no' means rejection.

In more detail, and as an example of the many possible ways scoring can be done, columns are added to each of the above questions representing the different risk levels. Against each question in each column, there is then an indicator of *mandatory*, *desirable*, *additional*, or *not applicable*, as in the table below with just three questions:

	Low risk	Medium risk	High risk
8. Are there visual or vibration alternatives to warning sounds?	Not applicable	Additional	Desirable
38. Has it been validated by an appropriate group of specialised professionals, health organisation or scientific society?	Additional	Desirable	Mandatory
51. Does the application logo relate to the purpose of the app	Not applicable	Additional	Additional

Confirming the answer yes to a question then either keeps the app in play if the indicator is *mandatory* (no would result in rejection), or scores 6 for *desirable* or an extra 4 (making 10 in total) for *additional*. A no to any *desirable* or *additional* question scores zero, as also does any answer where the risk level indicates *not applicable*.

So in the table above, if the app being assessed is high risk and the answer to Q38 is "No", then it is rejected immediately. If however it is medium risk, it scores 6, and low risk it scores 4.

The total score for each section is then divided by the number of scored questions to give an overall score. Scores below a set level result in rejection of the app. There are endless versions of this possible. One option to consider is giving higher weighting for some questions & lower weighting for others – thus in the examples above, Q38 might be given a higher weight than Q51.

An example of this is the Catalonian system further explained in annex 1. Also, MARS methodology uses full scoring.

Pass/fail

Involves questions being responded to electronically by the app developers, who also provide the evidence to support their answers. These can then then be audited. Scores, if they exist are not public.

Certification

For either of the above schemes, if an EU organisation were established to certify this scoring, certification/formal approval would be possible. Either public or private bodies could be envisaged to carry out third party certification based on the criteria outlined in the guidelines.

Some private initiatives already exist, such as Medappcare and Ourmobilehealth.

A list

Other options involve simply using the list as a set of questions to run through when considering commissioning, developing or acquiring an app. Such lists might be appropriately simplified for patients & carers, and perhaps for commissioners.

SOURCES

mHealth Evaluation Criteria

A separate document, attached as Annex 1, describes the resources consulted for this first draft. There are a substantial number of publications that have identified key evaluation criteria, so this aspect of the evaluation is well supported.

Many schemes are still in the process of being developed – some are happy to share their work in progress, others are not. Of those happy to share, the most complex is Catalonia's, comprising some 120 separate questions, with three different risk levels. This aims to produce as objective as possible an evaluation of apps, with some scoring to differentiate further. Not far behind in sophistication are published papers on the MARS and the mERA/WHO models. At a markedly less complex level are the checklists provided by come countries. Particularly notable among the simple checklists are ABACUS and KMNG.

Final choice will depend very much on who the Working Group seeks to target with the guidelines, whether anyone will be scoring them objectively and if so if there is a plan objectively to certify mHealth apps, on a voluntary basis.

GLOSSARY OF TERMS/ DEFINITIONS

Accessibility (of an interactive system): usability of a product, service, environment or facility by people with the widest range of capabilities

NOTE 1 The concept of accessibility addresses the full range of user capabilities and is not limited to users who are formally recognized as having a disability.

NOTE 2 The usability-orientated concept of accessibility aims to achieve levels of effectiveness, efficiency and satisfaction that are as high as possible considering the specified context of use, while paying particular attention to the full range of capabilities within the user population.

(ISO 9241-171:2008(E))

NB the UK mHealth apps classification has a term Equality which appears to be the same

Accountability: to be answerable for their actions and that there is redress when duties and commitments are not met (http://www.transparency-initiative.org)

Care pathway: a multidisciplinary outline of anticipated care, placed in an appropriate time frame, to help a patient with a specific condition or set of symptoms move progressively through a clinical experience to positive outcomes

Effectiveness: accuracy & completeness with which users achieve specific goals. (ISO 9241 11) or: extent to which planned activities are realized and planned results achieved. ((ISO 27000:2014)

Efficacy: a measure of the ability to produce the desired or intended result when operating in the care pathway for which it is designed. For patient-facing apps, the desired or intended result is a patient-relevant outcome

Efficiency: resources expended in relation to the accuracy & completeness with which users achieve specific goals. (ISO 9241-11)

Information security: preservation of confidentiality, integrity and availability of information

Confidentiality: property that information is not made available or disclosed to unauthorized individuals, entities, or processes

Integrity: property of accuracy and completeness

Availability: property of being accessible and usable upon demand by an authorized entity

(ISO 27000: 2014)

Interoperability: the ability of two or more systems or components to exchange information and to use the information that has been exchanged.

- Functional interoperability is the capability to reliably exchange information without error;

- Semantic interoperability is the ability to interpret, and, therefore, to make effective use of the information so exchanged.

(HL7)

mHealth App: a self-contained program or piece of software designed to fulfil a particular health or care-related purpose; an application, especially as downloaded by a user to a mobile platform

- health data can include fitness data if gathered for a medical purpose

- does not include off-label uses

- can be standalone or part of a service

Open data: Open data and content can be freely used, modified, and shared by anyone for any purpose. (http://opendefinition.org/)

Reliability; the ability of an app to yield the same result on repeated trials. (Also: property of consistent intended behaviour and results (ISO 27000:2008)

Safety: an unexpected problem or malfunction that may affect a patient's health or cause or contribute to an injury, for example a blood glucose meter giving an incorrect blood glucose reading, leading to incorrect treatment. (adapted from Health Products Regulatory Authority https://www.hpra.ie/homepage/medical-devices/safety-information). Brenda Reginatto

Technical stability: a measure of whether the app starts up reliably and completes its task without crashing

Transparency: managing and publishing information so that it is relevant and accessible and timely and accurate (http;//www.transparency-initiative.org)

Usability: The extent to which a product can be used by specified users to achieve specified goals with effectiveness, efficiency and satisfaction in a specified context of use. (ISO 9241¬11)

Validity: how well an app measures what it is supposed to measure. D. Focus

In priority order, the guidelines should deliver:

- Practical support or guidance

Terminology (clarity on use of terms)

- Legal (clarity)
- Organisational or procedural approach
- Criteria (to be used in assessing app)

LIST OF ANNEXES

1. Stakeholder Analysis – An overview of mHealth Evaluation Criteria

Stakeholder Analysis: An Overview of mHealth App Evaluation Criteria

Purpose

The purpose of this document is to summarise the evaluation criteria proposed by governments in the EU and in learned academic papers, with the objective of helping the Working Group decide on which criteria to select to produce guidelines for.

Those shown in this document were provided as a result of the request made at the initial WG meeting in Brussels on March 8th. No translations of non-English documents were immediately available so these have only been incorporated where it was obvious what the criteria chosen were.

Summary spreadsheet

The spreadsheet overleaf is an attempt to try to summarise the evaluation criteria that emerge from each of the sets of criteria that follow, purely in order to try to help the Working Group reach consensus. Of necessity it is subjective.

The criteria have been grouped together, again subjectively, to try to suggest some possible overall criteria for the Working Group to consider with the objective of attempting to establish criteria that are as independent as possible of each other.

Next steps

These criteria have now been used to produce a first draft of the app assessment guidelines, to which this annex is now attached.

	Appsalu	Cataloni	UK	Riezebo	Agarwal	MARS	DHAfB	ABACUS	KMNG	PAS277	
Relevance	✓										
Accessibility/equality	✓	✓	✓		✓						
Design	✓	✓		✓		✓		✓			**Usable & desirable** - a readily accessible & usable app that is desirable to use, easy to learn, appropriate for the task and interactive (could split into two or more)
Usability	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	
Training					✓				✓		
Appropriateness	✓			✓		✓					
User feedback/satisfaction					✓						
Entertainment/desirability						✓				✓	
Interaction		✓				✓					
Accuracy of description		✓				✓		✓			
Authors/developers - credibility/reputation	✓	✓		✓		✓		✓	✓		**Credible** - an app that is written & developed by well-regarded people/organisations, using regularly updated & well-evidenced content.
Updates & revisions to information	✓			✓				✓	✓	✓	
Content & information sources	✓✓	✓	✓	✓	✓	✓		✓	✓		
Standardisation											**Interoperable & transparent** - an app that uses widely-accepted standards & interoperates across platofrms tranparently and with other relevant apps
Interoperability			✓		✓					✓	
Open data			✓								
Transparency	✓		✓					✓			
Reliability/replicability		✓			✓	✓		✓	✓	✓	**Reliability & technical stability** - an app that can be relied on to repeat the same answer with the same inputs and remains technically stable over long periods
Technical support/stability	✓	✓	✓			✓		✓			
Bandwidth	✓									✓	
Scale limitations					✓					✓	
Risk management/safety	✓	✓	✓							✓	**Safe** - an app that does not expose the user to significant risk of harm
Effectiveness			✓	✓	✓	✓		✓			**Effective** - an app that does what it says on the tin, and delivers substantial benefit for low cost
Privacy & data protection	✓	✓	✓	✓	✓			✓	✓	✓	**Private & secure** - an app that explicitly states with whom data is shared and does so in an appropriately secure manner (could be two, separate)
Software security	✓	✓	✓	✓	✓				✓	✓	
eCommerce	✓							✓			
Advertising	✓							✓			
Data handling of the app				✓							
Infrastructure					✓						
Platform					✓						
Intervention delivery					✓						
Contextual adaptability					✓						
Regulatory compliance mechanism					✓					✓	
Adaptability of technology		✓									

Appsalud – the Andalucian approach

http://www.calidadappsalud.com/

Translation from: http://www.calidadappsalud.com/listado-completo-recomendaciones-app-salud/

Relevance & design

- Relevance
- Accessibility
- Design
- Usability

Quality & security of information

- Appropriateness (adecuación a la audiencia)
- Transparency
- Authorship
- Updates & revisions to information
- Content & information sources
- Risk management

Provision of services

- Technical support
- eCommerce
- Bandwidth
- Advertising

Privacy & security

- Privacy & data protection
- Software security

Catalonia

The Catalonian process essentially involves four stages. The first, “initial validation” is a simple sifting validation (see below for the important part of it).

The **initial validation** and classification of the app includes:

1. Validation of the **classification** proposed by the app developer
2. Validation of **concordance** between the proposed **version** and available in the official markets
3. **Smoke test:**
   - Installing and uninstalling
   - Validation basic functional: the application does what it is expected to make
   - Validation UI (user interface) Basic: main navigation screens and validation of initial user interface.

Note “smoke test” involves checking installation, functionality, usability, removal etc.

The next step is a risk assessment (see below). The resulting risk assessments drive the level of attention that each of the detailed questions receive.

technological criteria (*registration and management of information*) VS impact

Impacte Vs Criteris Tecnològics	Risc Baix	Risc Moderat	Risc Alt
Risc Baix	nivell 1	nivell 2	nivell 3
Risc Moderat	nivell 1	nivell 3	nivell 3 + auditoria de segureat
Risc Alt	nivell 2	nivell 3	nivell 3 + auditoria de segureat

criteria of content (*information provided*) VS impact

Impacte Vs Criteris Continguts	Risc Baix	Risc Moderat	Risc Alt
Risc Baix	nivell 1	nivell 1	nivell 2
Risc Moderat	nivell 1	nivell 2	nivell 3
Risc Alt	nivell 2	nivell 2	nivell 3

Columns are added to each of the detailed Catalonian questions representing the different risk levels. Against each question in each column, there is then an indicator of *mandatory, desirable, additional,* or *not applicable*, as in this table with just three questions:

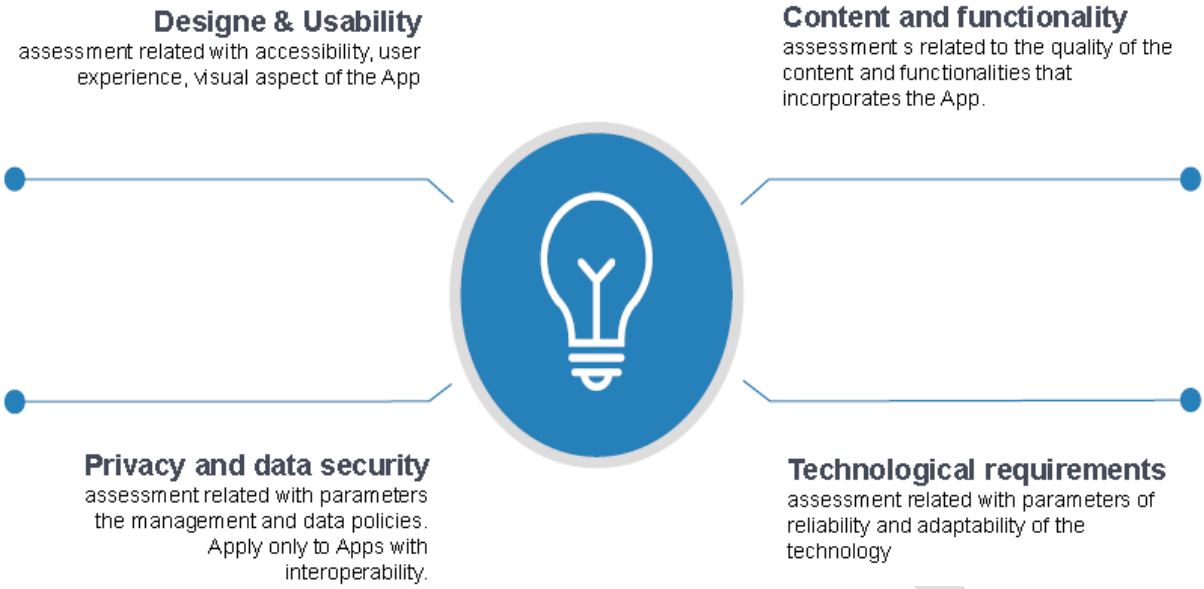
	Low risk	Medium risk	High risk
Question A	Not applicable	Additional	Desirable
Question B	Additional	Desirable	Mandatory
Question C	Not applicable	Additional	Additional

Confirming the answer yes to a question then either keeps the app in play if the indicator is *mandatory* (no would result in rejection), or scores 6 for *desirable* or an extra 4 (making 10 in total) for *additional*. A no to any *desirable* or *additional* question scores zero, as also does any answer where the risk level indicates *not applicable*.

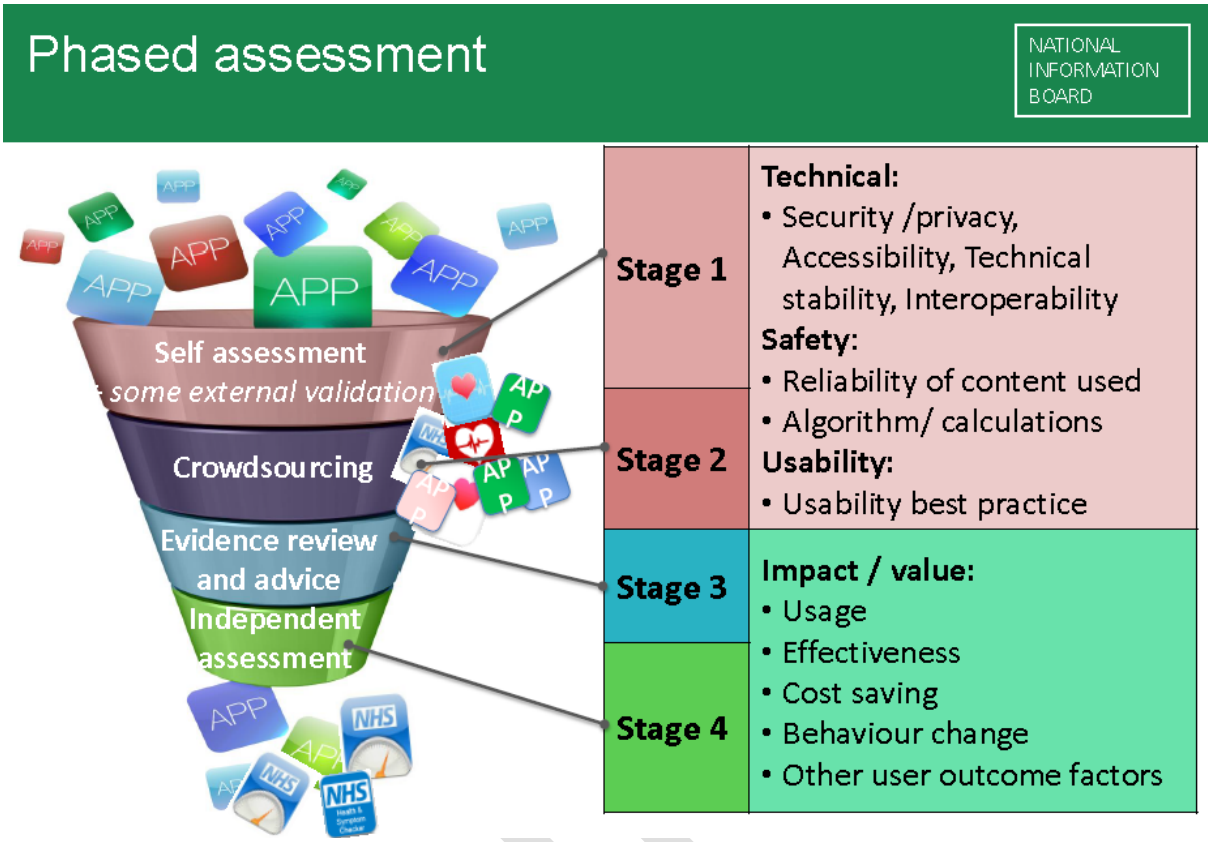
So in the table above, if the app being assessed is high risk and the answer to QB is “No”, then it is rejected immediately. If however it is medium risk, it scores 6, and low risk it scores an additional 4.

The total score for each section is then divided by the number of scored questions to give a figure between 5 & 10. Scores of less than 5 result in rejection of the app.

The detailed questions fit into one of four categories:



The UK Approach



Stage 1 – self-assessment questions

YES

Privacy and Security

Are you looking after information in line with the Data Protection Act and protecting personal data in line with best practice?

YES

Safety

What risk could your App pose and how is that risk communicated and managed?

YES

Interoperability

When your App integrates with other clinical systems what standards are being met and is best practice being followed?

YES

Open Data

Contributing to open data benefits the health community. Meeting best practice raises the profile and usefulness of your App.

YES

Equality

What efforts have been made to make your App available to and usable by the widest number of users?

YES

Effectiveness

What efforts and evidence can you provide to support the effectiveness of your App?

YES

Transparency

Being honest and open about who makes, publishes and benefits from the users usage of the App builds trust and encourages uptake.

YES

Usability

What rigour has been applied to making the App free of technical and usability issues?

YES

Registration

Information about registration form

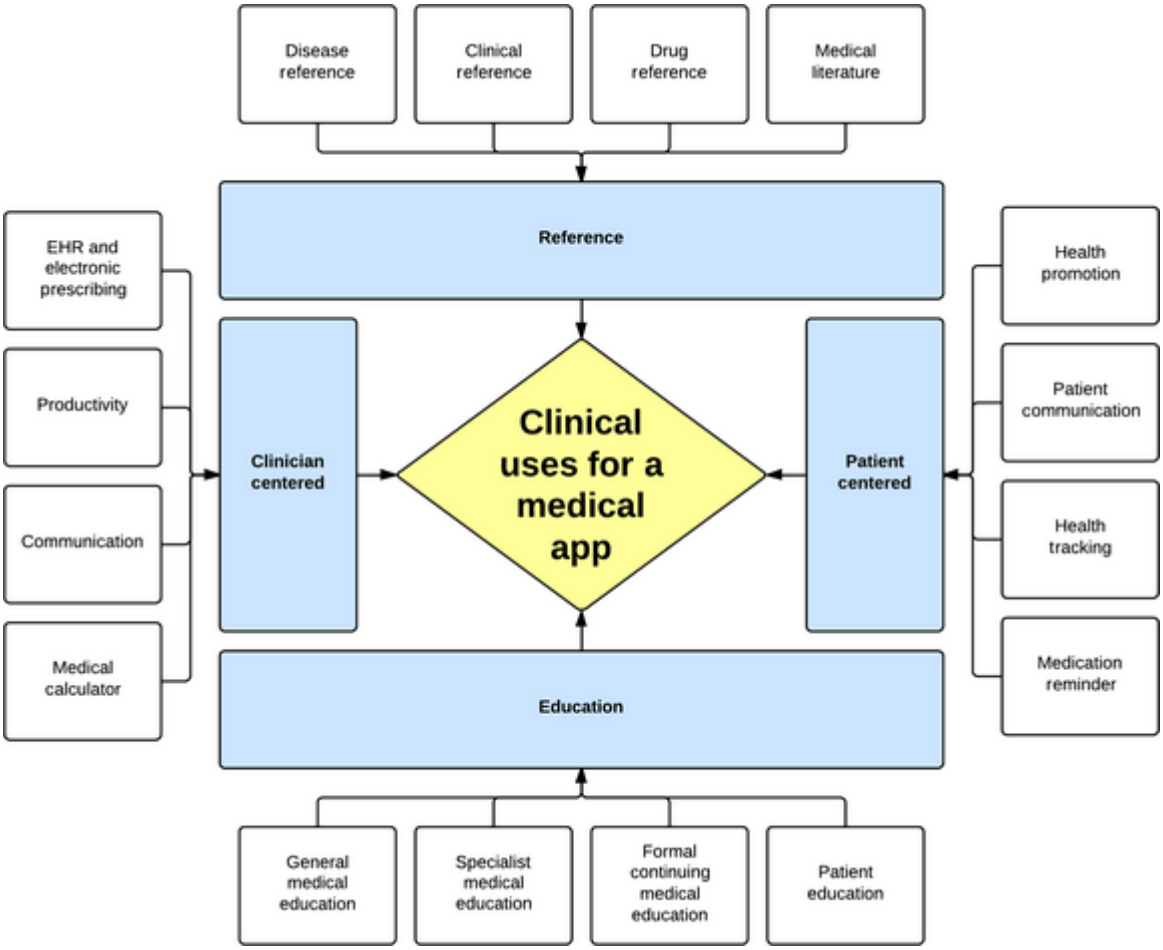
YES

Technical Stability

info about tech stability

How to identify, assess and utilise mobile medical applications in clinical practice **T**. D. Aungst et al, International Journal of Clinical Practice > Vol 68 Issue 2first published online: 26 Jan 2014 http://onlinelibrary.wiley.com/doi/10.1111/ijcp.12375/full

Figure 1



Peer-reviewing of mHealth applications, R. J. Riezebos Oct 2014 (PhD thesis, Univ of Amsterdam) http://dare.uva.nl/cgi/arno/show.cgi?fid=573074

Pages 9-14 are especially relevant as a pick-list of potential criteria

P24 has a key diagram from a survey of 135 consumers:

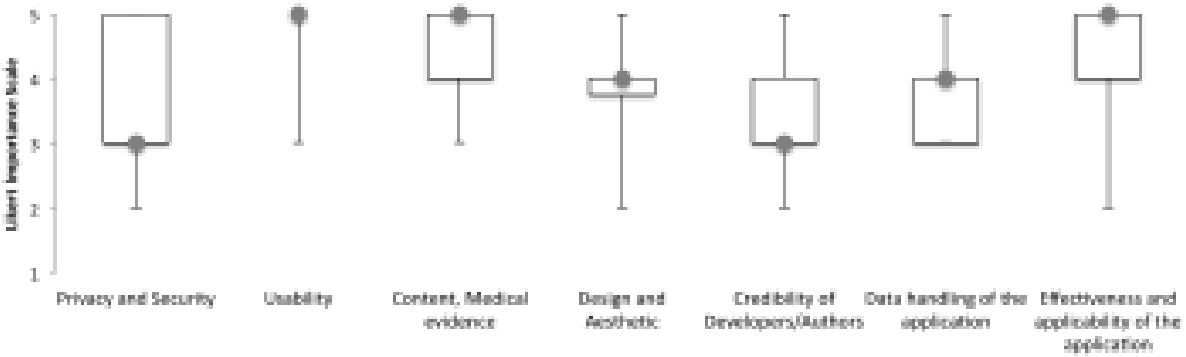


Figure 6 - Identification of important review elements for developers
Likert Importance scale: 1= Unimportant – 5=Very important, • = median, box = 25-75% interquartile range, line = range of responses

From this, the criteria in a rough order of apparent importance are:

1. Usability
2. Content, medical evidence
3. Effectiveness & applicability of the application
4. Design & aesthetic
5. Privacy & security
6. Data handling of the app
7. Credibility of developers/authors

Guidelines for reporting of health interventions using mobile phones: mobile health (mHealth) evidence reporting and assessment (mERA) checklist. Agarwal S, et al BMJ 2016; 352:i1174

"mERA was developed as a checklist of items which could be applied by authors developing manuscripts that aim to report on the effectiveness of mHealth interventions and by peer reviewers and journal editors reviewing such evidence. mERA aims to provide guidance for complete and transparent reporting on studies evaluating and reporting on the feasibility and effectiveness of mHealth interventions…" (*Scope* paragraph)

Criteria	**Item no**	**Notes**
Infrastructure (population level)	1	Clearly presents the availability of infrastructure to support technology operations in the study location. This refers to physical infrastructure such as electricity, access to power, connectivity etc. in the local context. Reporting X% network coverage rate in the country is insufficient if the study is not being conducted at the country level
Technology platform	2	Describes and provides justification for the technology architecture. This includes a description of software and hardware and details of any modifications made to publicly available software
Interoperability/Health information systems (HIS) context	3	Describes how mHealth intervention can integrate into existing health information systems. Refers to whether the potential of technical and structural integration into existing HIS or programme has been described irrespective of whether such integration has been achieved by the existing system
Intervention delivery	4	The delivery of the mHealth intervention is clearly described. This should include frequency of mobile communication, mode of delivery of intervention (that is, SMS, face to face, interactive voice response), timing and duration over which delivery occurred
Intervention content	5	Details of the content of the intervention are described. Source and any modifications of the intervention content is described
Usability/content testing	6	Describe formative research and/or content and/or usability testing with target group(s) clearly identified, as appropriate
User feedback	7	Describes user feedback about the intervention or user satisfaction with the intervention. User feedback could include user opinions about content or user interface, their perceptions about usability, access, connectivity, etc.
Access of individual participants	8	Mentions barriers or facilitators to the adoption of the intervention among study participants. Relates to individual-level structural,

Criteria	Item no	Notes
		economic and social barriers or facilitators to access such as affordability, and other factors that may limit a user's ability to adopt the intervention
Cost assessment	9	Presents basic costs assessment of the mHealth intervention from varying perspectives. This criterion broadly refers to the reporting of some cost considerations for the mHealth intervention in lieu of a full economic analysis. If a formal economic evaluation has been undertaken, it should be mentioned with appropriate references. Separate reporting criterion are available to guide economic reporting
Adoption inputs/ programme entry	10	Describes how people are informed about the programme including training, if relevant. Includes description of promotional activities and/or training required to implement the mHealth solution among the user population of interest
Limitations for delivery at scale	11	Clearly presents mHealth solution limitations for delivery at scale
Contextual adaptability	12	Describes the adaptation, or not, of the solution to a different language, different population or context. Any tailoring or modification of the intervention that resulted from pilot testing/usability assessment is described
Replicability	13	Detailed intervention to support replicability. Clearly presents the source code/screenshots/ flowcharts of the algorithms or examples of messages to support replicability of the mHealth solution in another setting
Data security	14	Describes the data security procedures/ confidentiality protocols
Compliance with national guidelines or regulatory statutes	15	Mechanism used to assure that content or other guidance/information provided by the intervention is in alignment with existing national/regulatory guidelines and is described
Fidelity of the intervention	16	Was the intervention delivered as planned? Describe the strategies employed to assess the fidelity of the intervention. This may include assessment of participant engagement, use of backend data to track message delivery and other technological challenges in the delivery of the intervention

"The mERA checklist was developed by a group of experts assembled as part of the WHO mTERG, reflecting a diversity of geographical, gender, and domain expertise. Contributors outside of mTERG were recruited through professional and academic networks; their representation could have been biased towards experts focused on public health interventions in low and middle income country programmes."

MARS

Mobile App Rating Scale: A New Tool for Assessing the Quality of Health Mobile Apps, Stoyanov SR et al, JMIR mHealth uHealth 2015;3(1): e27 http://mhealth.jmir.org/2015/1/e27/

"The objective of this study was to develop a reliable, multidimensional measure for trialling, classifying, and rating the quality of mobile health apps." Text subsequently refers to "users, health professionals, and researchers" as target for MARS.

App Quality Ratings

The Rating scale assesses app quality on four dimensions. All items are rated on a 5-point scale from "1. Inadequate" to "5. Excellent". Circle the number that most accurately represents the quality of the app component you are rating. Please use the descriptors provided for each response category.

SECTION A

Engagement – fun, interesting, customisable, interactive (e.g. sends alerts, messages, reminders, feedback, enables sharing), well-targeted to audience

1**. Entertainment**: Is the app fun/entertaining to use? Does it use any strategies to increase engagement through entertainment (e.g. through gamification)?

1 Dull, not fun or entertaining at all
2 Mostly boring
3 OK, fun enough to entertain user for a brief time (< 5 minutes)
4 Moderately fun and entertaining, would entertain user for some time (5-10 minutes total)
5 Highly entertaining and fun, would stimulate repeat use

2. **Interest**: Is the app interesting to use? Does it use any strategies to increase engagement by presenting its content in an interesting way?

1 Not interesting at all
2 Mostly uninteresting
3 OK, neither interesting nor uninteresting; would engage user for a brief time (< 5 minutes)
4 Moderately interesting; would engage user for some time (5-10 minutes total)
5 Very interesting, would engage user in repeat use

3. **Customisation**: Does it provide/retain all necessary settings/preferences for apps features (e.g. sound, content, notifications, etc.)?

1 Does not allow any customisation or requires setting to be input every time
2 Allows insufficient customisation limiting functions
3 Allows basic customisation to function adequately
4 Allows numerous options for customisation
5 Allows complete tailoring to the individual's characteristics/preferences, retains all settings

4. **Interactivity**: Does it allow user input, provide feedback, contain prompts (reminders, sharing options, notifications, etc.)? Note: these functions need to be customisable and not overwhelming in order to be perfect.

1 No interactive features and/or no response to user interaction
2 Insufficient interactivity, or feedback, or user input options, limiting functions
3 Basic interactive features to function adequately
4 Offers a variety of interactive features/feedback/user input options
5 Very high level of responsiveness through interactive features/feedback/user input options

5. **Target group**: Is the app content (visual information, language, design) appropriate for your target audience?

1 Completely inappropriate/unclear/confusing
2 Mostly inappropriate/unclear/confusing
3 Acceptable but not targeted. May be inappropriate/unclear/confusing
4 Well-targeted, with negligible issues
5 Perfectly targeted, no issues found

SECTION B

Functionality – app functioning, easy to learn, navigation, flow logic, and gestural design of app

6. **Performance**: How accurately/fast do the app features (functions) and components (buttons/menus) work?

1 App is broken; no/insufficient/inaccurate response (e.g. crashes/bugs/broken features, etc.)
2 Some functions work, but lagging or contains major technical problems
3 App works overall. Some technical problems need fixing/Slow at times
4 Mostly functional with minor/negligible problems
5 Perfect/timely response; no technical bugs found/contains a 'loading time left' indicator

7. **Ease of use**: How easy is it to learn how to use the app; how clear are the menu labels/icons and instructions?

1 No/limited instructions; menu labels/icons are confusing; complicated
2 Useable after a lot of time/effort
3 Useable after some time/effort
4 Easy to learn how to use the app (or has clear instructions)
5 Able to use app immediately; intuitive; simple

8. **Navigation**: Is moving between screens logical/accurate/appropriate/ uninterrupted; are all necessary screen links present?

1 Different sections within the app seem logically disconnected and random/confusing/ navigation is difficult
2 Usable after a lot of time/effort
3 Usable after some time/effort
4 Easy to use or missing a negligible link
5 Perfectly logical, easy, clear and intuitive screen flow throughout, or offers shortcuts

9. **Gestural design**: Are interactions (taps/swipes/pinches/scrolls) consistent and intuitive across all components/screens?

1 Completely inconsistent/confusing
2 Often inconsistent/confusing
3 OK with some inconsistencies/confusing elements
4 Mostly consistent/intuitive with negligible problems
5 Perfectly consistent and intuitive

SECTION C

Aesthetics – graphic design, overall visual appeal, colour scheme, and stylistic consistency

10. **Layout**: Is arrangement and size of buttons/icons/menus/content on the screen appropriate or

zoomable if needed?

1 Very bad design, cluttered, some options impossible to select/locate/see/read device display
not optimised
2 Bad design, random, unclear, some options difficult to select/locate/see/read
3 Satisfactory, few problems with selecting/locating/seeing/reading items or with minor screensize problems
4 Mostly clear, able to select/locate/see/read items
5 Professional, simple, clear, orderly, logically organised, device display optimised. Every design component has a purpose

11. **Graphics**: How high is the quality/resolution of graphics used for buttons/icons/menus/content?

1 Graphics appear amateur, very poor visual design - disproportionate, completely stylistically inconsistent
2 Low quality/low resolution graphics; low quality visual design – disproportionate, stylistically inconsistent
3 Moderate quality graphics and visual design (generally consistent in style)
4 High quality/resolution graphics and visual design – mostly proportionate, stylistically consistent
5 Very high quality/resolution graphics and visual design - proportionate, stylistically consistent throughout

12. **Visual appeal**: How good does the app look?

1 No visual appeal, unpleasant to look at, poorly designed, clashing/mismatched colours
2 Little visual appeal – poorly designed, bad use of colour, visually boring
3 Some visual appeal – average, neither pleasant, nor unpleasant
4 High level of visual appeal – seamless graphics – consistent and professionally designed
5 As above + very attractive, memorable, stands out; use of colour enhances app features/menus

SECTION D

Information – Contains high quality information (e.g. text, feedback, measures, references) from a credible source. Select N/A if the app component is irrelevant.

13. Accuracy of app description (in app store): Does app contain what is described?

1 Misleading. App does not contain the described components/functions. Or has no description
2 Inaccurate. App contains very few of the described components/functions
3 OK. App contains some of the described components/functions
4 Accurate. App contains most of the described components/functions
5 Highly accurate description of the app components/functions

14. **Goals**: Does app have specific, measurable and achievable goals (specified in app store description or within the app itself)?

N/A Description does not list goals, or app goals are irrelevant to research goal (e.g. using a game for educational purposes)
1 App has no chance of achieving its stated goals
2 Description lists some goals, but app has very little chance of achieving them
3 OK. App has clear goals, which may be achievable.

4 App has clearly specified goals, which are measurable and achievable
5 App has specific and measurable goals, which are highly likely to be achieved

15**. Quality of information**: Is app content correct, well written, and relevant to the goal/topic of the app?

N/A There is no information within the app
1 Irrelevant/inappropriate/incoherent/incorrect
2 Poor. Barely relevant/appropriate/coherent/may be incorrect
3 Moderately relevant/appropriate/coherent/and appears correct
4 Relevant/appropriate/coherent/correct
5 Highly relevant, appropriate, coherent, and correct

16. **Quantity of information**: Is the extent coverage within the scope of the app; and comprehensive but concise?

N/A There is no information within the app
1 Minimal or overwhelming
2 Insufficient or possibly overwhelming
3 OK but not comprehensive or concise
4 Offers a broad range of information, has some gaps or unnecessary detail; or has no links to
more information and resources
5 Comprehensive and concise; contains links to more information and resources

17. **Visual information**: Is visual explanation of concepts – through charts/graphs/images/videos, etc.
– clear, logical, correct?

N/A There is no visual information within the app (e.g. it only contains audio, or text)
1 Completely unclear/confusing/wrong or necessary but missing
2 Mostly unclear/confusing/wrong
3 OK but often unclear/confusing/wrong
4 Mostly clear/logical/correct with negligible issues
5 Perfectly clear/logical/correct

18. **Credibility**: Does the app come from a legitimate source (specified in app store description or within the app itself)?

1 Source identified but legitimacy/trustworthiness of source is questionable (e.g. commercial
business with vested interest)
2 Appears to come from a legitimate source, but it cannot be verified (e.g. has no webpage)
3 Developed by small NGO/institution (hospital/centre, etc.) /specialised commercial business, funding body
4 Developed by government, university or as above but larger in scale
5 Developed using nationally competitive government or research funding (e.g. Australian Research Council, NHMRC)

19. **Evidence base**: Has the app been trialled/tested; must be verified by evidence (in published scientific literature)?

N/A The app has not been trialled/tested

1 The evidence suggests the app does not work
2 App has been trialled (e.g., acceptability, usability, satisfaction ratings) and has partially positive outcomes in studies that are not randomised controlled trials (RCTs), or there is little or no contradictory evidence.
3 App has been trialled (e.g., acceptability, usability, satisfaction ratings) and has positive outcomes in studies that are not RCTs, and there is no contradictory evidence.
4 App has been trialled and outcome tested in 1-2 RCTs indicating positive results
5 App has been trialled and outcome tested in > 3 high quality RCTs indicating positive results

SECTION E
App subjective quality

20. Would you recommend this app to people who might benefit from it?
1 Not at all I would not recommend this app to anyone
2 There are very few people I would recommend this app to
3 Maybe There are several people whom I would recommend it to
4 There are many people I would recommend this app to
5 Definitely I would recommend this app to everyone

21. How many times do you think you would use this app in the next 12 months if it was relevant to you?
1 None
2 1-2
3 3-10
4 10-50
5 >50

22. Would you pay for this app?
1 No
3 Maybe
5 Yes

23. What is your overall star rating of the app?
1 « One of the worst apps I've used
2 ««
3 ««« Average
4 «««««
5 «««««« One of the best apps I've used

Scoring
App quality scores for SECTION
A: Engagement Mean Score = _____
B: Functionality Mean Score = _____
C: Aesthetics Mean Score = _____
D: Information Mean Score = _____
App quality mean Score = _____
App subjective quality Score = _____
App-specific

These added items can be adjusted and used to assess the perceived impact of the app on the user's knowledge, attitudes, intentions to change as well as the likelihood of actual change in the target health behaviour.

SECTION F

1. **Awareness**: This app is likely to increase awareness of the importance of addressing [insert target health behaviour]

Strongly disagree 1 2 3 4 5 Strongly Agree

2. **Knowledge**: This app is likely to increase knowledge/understanding of [insert target health behaviour]

<scoring as above>

3. **Attitudes**: This app is likely to change attitudes toward improving [insert target health behaviour]

<scoring as above>

4. **Intention to change**: This app is likely to increase intentions/motivation to address [insert target health behaviour]

<scoring as above>

5. **Help seeking**: Use of this app is likely to encourage further help seeking for [insert target health behaviour] (if it's required)

<scoring as above>

6. **Behaviour change**: Use of this app is likely increase/decrease [insert target health behaviour] Strongly disagree

<scoring as above>

Digital-HealthAnwendungen für Bürger (Digital health apps for citizens)

https://www.bertelsmann-stiftung.de/fileadmin/files/BSt/Publikationen/GrauePublikationen/Studie_VV_Digital-Health-Anwendungen_2016.pdf

Paper contains much detailed classification information and research.

One classification:

Application

- Health behaviour stage (?Schritte des Gesundheitshandelns) – eg information, assessment, monitoring, behaviour change motivation etc.
- Function
- User

Target group

- Health status
- Age group
- Gender

Usage context

- Application/theme
- Level of care
- Performance level (?Leistungssektor)

Technology

Provider

Quality & usability

Authority

ABACUS

http://libguides.library.arizona.edu/c.php?g=122854&p=802639

This excellent short checklist is an acronym of the question headings:

Accuracy, Bias/objectivity, Authority, Currency/timeliness, Usability, Scope/completeness,

Excellent for a non-expert user. The full set of questions is:

Accuracy

- Is the medical information contained in the app based on sound medical research and evidence? Can the information in the app be verified by another source?
- Are there references/sources included so that you can verify the information? Are these references reliable? (For example, a citation to a drug company website does not have the same weight as an article from*JAMA*.)
- Are there grammatical and spelling errors? (This may be a "tell" - if the information isn't even spelled correctly, maybe the information itself isn't correct.)
- Does the app do what it intends to do? Is there any potential for patient harm?

Bias/Objectivity

- Is the information showing just one point of view or is it sponsored by a company that is trying to sell something?
- What kind of organization sponsored the app? A pharmaceutical company? A non-profit organization? A reputable journal?
- Is advertising clearly marked and distinguishable from the informational/medical content? Can you tell if the information you are reading is advertisement?
- Does the app use data improperly to promote a position or a product, or is it unbiased/neutral?

Authority

- Who developed the app? What are the person's or sponsoring organization's credentials? Are they an expert in the content presented in the app? What do you know about them?
- Is the person backed by a known organization? (Be careful here... some "organizations" may simply be unreliable groups operating out of someone's basement; try to go with authoritative sources, like the National Library of Medicine.)

- Do experts review the content provided in the app and are these "experts" real authorities on the content?
- Can you easily find contact information in the app or on its download/information page? Check the about us link/seller information, usually found on the app's download page. What is the purpose of the organization? Is it trying to sell something or is it an unbiased, peer-reviewed information source?

Currency/Timeliness

- When was the app created and/or last updated?
- Does the app provide regular updates when new content or technological upgrades are required?
- Has there been more recent research on the content in the app? Many medical treatments change with the publication of new studies. What was published a year ago may be outdated now.

Usability

- Does the app work reliably and stably on the device you are using?
- Is navigation smooth and intuitive?
- Is the app efficient and effective? For example, is the type of content usable on a small screen (e.g., radiological images)? Is data entry easy?
- Is the app appropriate for the target audience (e.g., patient info apps are in plain language)?
- Does the app author provide technical support for the app?
- Is the app stand-alone (meaning you can use it without a wi-fi or Internet connection)? This is just a good thing to note so you are aware about whether the app can be used without an Internet connection.

Scope/Completeness

- Is the medical information presented in the app complete?
- Are there sources given for additional information?
- Who is the target audience - is the app targeted for use by medical professionals, patients, others?

Be sure to ask yourself:

- Why did the person/organization create the app?
- What's in it for them or are they trying to sell me something?
- Is the creator of the app an expert in the content presented in the app?
- Can I verify the information being presented to me in the app and is it accurate?
- Is there a way I can contact the app developer to provide feedback or ask a question?
- Are there any login requirements or privacy issues that I need to know about if I choose to use this app? Will my use of this app be tracked in any way?
- Is there a disclaimer that states any impact on clinical decision making, patient safety?

***TEST** before you use:*

- TEST the app before you use it in clinical care - create clinical scenarios and test.
- As you test, observe and evaluate the app according to the above ABACUS framework. Does it pass the Accurate, un-Biased, Authoritative, Current, Usable, Scope/Completeness benchmarks in multiple case scenarios?

KMNP App checker

file:///C:/Users/Charles/Downloads/KNMG_MedischeApp_170x240_EN.pdf

A nice short document with 19 pass/fail questions (some multi-part)

Headings: CE Mark, Functionality, Content quality and (clinical) relevance, Ese of use, Privacy, Security, Final evaluation

Well worth checking out.

BSI PAS 277

http://shop.bsigroup.com/forms/PASs/PAS-2772015/ (it is free to download)

The quality criteria proposed (page 5) are:

Regulatory and legal compliance – it is important to be aware of specific regulations and laws that might apply, and to ensure that compliance can be described in terms used in the regulations.

Functionality – covers the functions that are required to support the intended use of the app for the user, and functions the app requires to meet the relevant needs of any other stakeholders.

User & user experience – including considerations of accessibility for different types of users, and how using the app might fit in with related activities that the user performs. Reliability, performance & scalability.

Reliability, performance, and scalability – to cover both the performance of the app itself, and the supporting infrastructure, such as web services that the app may rely on.

Security & privacy – to include effective controls over the app and information that it collects, while ensuring that before choosing to use the app, the user is made aware of how personal information is collected, stored and used.

Safety – incl. patient safety where relevant, as well as safety considerations that would apply to any software product.

Compatibility & portability – including compatibility of the app with different platform configurations and the ways that information collected or used by the app may be reused, under appropriate privacy controls.

Maintainability – it is important that the app is maintained so that it can deliver the intended use, or at least until support is discontinued by the app publisher. It should cover all the considerations that are relevant to the reliable and cost effective provision of maintenance services and configuration control.

Other relevant documentation

https://www.rcplondon.ac.uk/file/175/download?token=5nTJceC1 – Guidance from the UK Royal College of Physicians, primarily concerned with whether an app is a medical device.

http://www.fda.gov/downloads/MedicalDevices/.../UCM263366.pdf - Guidance for industry & FDA staff

http://www.fda.gov/downloads/medicaldevices/deviceregulationandguidance/guidancedocuments/ucm429674.pdf - FDA guidance for low risk devices

http://www.raeng.org.uk/publications/reports/health-apps-regulation-and-quality-control - a summary of an AMedSci event in late 2014

http://www.jmir.org/2012/5/e128/ (Odzalga et al 2012) – "The Smartphone in Medicine: A Review of Current and Potential Use…"

http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2284448 (Cortez 2013) – "The mobile health revolution?"

http://mhealth.jmir.org/2016/2/e34/#Abstract (Weda et al 2016) – "Characterization of Apps and Other e-Tools for Medication Use: Insights Into Possible Benefits and Risks"

http://health.gov/healthliteracyonline/2010/Web_Guide_Health_Lit_Online.pdf - a health literacy guide

There's an interesting Peer-review (JMIR) at http://tinyurl.com/appsform Here is a list of other app comparison sites that, with one exception (in German) did not when accessed reveal their criteria or detailed questions:

UK
https://orcahealth.com/
Germany
http://www.appcheck.de/
https://www.healthon.de/de
France
http://www.dmd-sante.com/
USA
http://www.zurinstitute.com/mentalhealthapps_resources.html
http://www.eatrightpro.org/resources/media/trends-and-reviews/app-reviews
https://www.happtique.com/home/
http://www.imedicalapps.com/about/
http://diabetes.ufl.edu/my-diabetes/diabetes-resources/diabetes-apps/
In addition, the following are sites of Working Group members
http://myhealthapps.net/
http://www.ourmobilehealth.co.uk/our-services.html
http://www.medappcare.com/en/

Here are some security references:

https://resilience.enisa.europa.eu/cloud-security-and-resilience/publications/cloud-computing-benefits-risks-and-recommendations-for-information-security
https://www.enisa.europa.eu/?came_from=https%253A//www.enisa.europa.eu/topics/iot-and-

smart-infrastructures/mobile-applications/smartphone-security-1/top-ten-risks
https://www.owasp.org/index.php/OWASP_Cloud_%E2%80%90_10/Initial_Pre-Alpha_List_of_OWASP_Cloud_Top_10_Security_Risks

http://www.iso.org/iso/fr/home/standards/management-standards/iso27001.htm

The best short paper on app risk assessment is:

mHealth and Mobile Medical Apps: A Framework to Assess Risk and Promote Safer Use
http://www.jmir.org/2014/9/e210/