Annex 1
Stakeholders' Positions on Assessment Criteria

Contents

1 Overview by criteria/stakeholder group 2

2 Meaning of criteria by stakeholder groups 3

2.1 Patients 3

2.2 Healthcare Professionals (HCP) 4

2.3 Industry 4

2.4 Public Authorities 8

2.5 Payers and Social Health Insurance 9

2.6 Research and Academia 10

1 Overview by criteria/stakeholder group

Question: Is this criterion relevant for assessing validity/reliability ?													
Answers for assessment of apps								***Answers for assessment of data***					
	Patients	Healthcare Professionals	Public Authorities	Payers Social Health Insurance	Research / Academia	Industry		Patients	Healthcare Professionals	Public Authorities	Payers Social Health Insurance	Research / Academia	Industry
Privacy	Y	Y	Y	Y	Y	Not answered							N
Transparency	Y	Y	Y	Y	Y	Not answered							N
Safety	Y	Y	Y	Y	Y	Not answered							N
Reliability	Y	Y	Y	Y	Y	Not answered							Y
Validity	Y	Y	Y	Y	Y	Not answered							Y
Interoperability	Y	Y	Y	Y	Y	Not answered							N
Technical stability	Y	Not answered	Y	Y	Y	Not answered							N
Effectiveness	Y	Y	Y: Glasgow; No: Infarmed and Andalucia	Y	Y	Not answered							N
Efficacy	Y	Not answered	Unsure / optional	Y	Y	Not answered							N
Efficiency	Y	Not answered	Unsure	Y	Y	Not answered							N
Accessibility	N - but see new criterion below	Y	Y with caveat (Glasgow)	Y	Y	Not answered				Not answered			N
Usability/desirability	N - but see new criterion below	Not answered	Y	Not answered	Y	Not answered							N
Scalability	N	Not answered	Not answered	Not answered	Y	Not answered							N
						Not answered							N
User experience, user centered design, suitability	Y	These 2 additional criteria are proposed by patient representatives, so other stakeholders have not had a chance to consider				Not answered							
Security	Y					Not answered							

2 Meaning of criteria by stakeholder groups

2.1 Patients

Privacy: App compliance with applicable laws and guidelines.

Whilst the privacy criteria should largely be covered by the Code of Conduct on Privacy and GDPR, the responder felt privacy should be flagged in the assessment questionnaire as a reminder of their legal obligations and best practice (where voluntary). It doesn't need to be an assessment criterion, just provide links to the Code, GDPR provisions and national data protection rules for further information.

Developers and manufacturers of devices using apps to collect and process data, should warn and inform mHealth apps and device users at all times and well in advance before their data are anonymized for use for purposes other than historical, statistical or scientific purposes. In such cases, users should be given the possibility, if they so wish, to delete any personal health or other data they might have entered in the app. The nature and type of information requested for recording in any given mHealth app should be commensurate with the purpose of the app and the requirements for optimal usability and functionality.

Transparency: Identification of the originators/sponsors of apps.

The user should be informed which of his/her personal data the app needs for its smooth function, how his/her data are handled after he registered them in the app, to whom and when they might be transmitted for further processing. What happens to his/her data if he/she stops using the app, will the app prompt the user to re-use the app or inform him that he may delete his data if he does not intend to re-use it. Moreover, the user should be informed if he/she can delete permanently any data he/she might have entered, if he/she does not agree with subsequently.

Safety: The need to avoid any harm to patients.

Reliability: Consistency in the functioning of an app.

Interoperability: The need for an app to function on multiple platforms.

Technical stability: The developer should provide a regularly updated FAQ in which solutions are offered for the most frequent problems reported by the users.

Effectiveness: In order to build user/patient/Health Care Professional trust in an app, evidence of studies, trials etc. where relevant should be referred to. As a minimum, the app needs to be evaluated against any claimed health benefit or improved health outcome.

Efficacy: Medical accuracy; patient willingness to record data using app (aka efficiency).

Usability/desirability: App eases of use for the widest group of user possible, taking into account, different physical or mental abilities or impairments (where relevant).

Two additional criteria were proposed

1. **User experience, user centred design, suitability.** User experience (UX) refers to a person's total experience using a particular product, system or service. It is defined in ISO 9241-210, (Ergonomics of human-system interaction) as "a person's perceptions

and responses that result from the use or anticipated use of a product, system or service". It refers to a person's holistic experience (physical, cognitive, emotive, beliefs, preferences, behaviours etc.) and as such is a multifaceted concept, which incorporates consideration of the user, the system and the context of use as a triad of important influencing factors. It is a consequence of brand image, presentation, functionality, system performance, interactive behaviour and assistive capabilities of the interactive system, the user's internal and physical state resulting from prior experiences, attitudes, skills and personality, and the context of use.

2. **Security**: Users expect that their personal health data to be stored in a secure way and that the mHealth solution does not require more information than needed.

2.2 Healthcare Professionals (HCP)

Privacy: The patient-HCP relationship is based on trust. Therefore, every measure has to be taken to ensure that personal health data is protected from unauthorised access or misuse.

Transparency: Providing HCP with relevant and trustworthy information about the app and the developer is crucial for reliable decision making processes on whether the app may/will be used in a health care context.

Safety: For HCPs, safety is an important criterion, not least of all due to liability concerns.

Reliability and Validity: Apps for medical purposes do not always respect established clinical guidelines or do not provide reliable implementations of algorithms.

Interoperability: this matters for HCPs, mainly when they discover that they cannot use desired data from not interoperable devices or data formats.

Technical stability: HCPs react negatively to technology they cannot rely on.

Efficiency: Reduction of the frequency of face-to-face visits for instance or reduction on medication, decrease on hospitalization rate.

Usability / Desirability: HCPs will use any tool that will support them fulfilling their needs. If they have a choice, they will choose the app that will give them the highest level of comfort. An excellent usability will be an advantage and is also important to prevent usage errors. High usability and desirability are also important in the context of app attrition, i.e. to ensure that users continue using the apps.

Scalability: Especially in professional (clinical) contexts of HCPs, scalability is an important aspect. For instance, there are obligatory documentation periods for many types of health data, potentially leading to a growing amount of data having to be handled.

2.3 Industry

Validity and reliability: Industry perspective is that the original mandate for the work concerned only validity and reliability of the data, as specified in the original mandate for the WG. Industry acknowledges that other criteria may be relevant, but suggests a stepwise approach starting off with data validity and reliability.

Reliability of data: stability, repeatability, or precision of the data. Definition provided by ISO/TR 18307:2001(en), 3.51

> Note 1: Does the app produce repeatable and reliable results i.e. does the app always give the same results for the same set of parameters?
>
> Note 2: The criterion refers to output data

Data validity: verification of correctness (reflecting the true situation) ISO/TR 18307:2001(en), 3.53

> Note 1: Are the results meaningful, do they have some proven scientific basis and value?
>
> Note 2: It refers to output data

Technical stability: The guidelines must avoid addressing the technology of an app. The technical stability of an app has little to do with whether the app produces reliable and valid data. This issue is more of a market based issued. If an app is technically unstable, it will likely not survive on the market place as it will lead to a lack of usage, poor reviews, and no downloads. If this exercise were to focus on the overall quality assessment of an app, then its underlying technical stability could be considered as a potential criterion.

Accessibility: The criterion here could vary widely. 'Accessibility' to us could mean it is 'accessible' to the consumer on the market place (i.e. it is available on app stores). Or 'accessibility' could mean that it adheres to national, EU and international laws and guidelines for those with disabilities or limited cognitive ability.

Privacy: This criterion includes complying with the requirements of national data protection laws and the principles laid down in the Data Protection Directive and the ePrivacy Directive. As of May 2018 there will be additional obligations on data protection found in the GDPR. The criterion for a proper assessment would need to cover a variety of issues including but not limited to:

- lawfulness of processing (consent, legitimate interest, performance of a contract, data minimisation, etc.)
- transparency
- information to be provided to data subjects
- rights of access, rectification, erasure, portability, objection
- automated decision making
- data protection by design & by default
- controller & processor obligations including liability
- record keeping
- data breach notification
- data protection impact assessment

- data protection officer designation

- international data transfers

Issues related to privacy have nothing to do with assessing the reliability/validity of the data that an mHealth app produces. An mHealth app can be 'non-compliant' with data privacy law, but still produce valid/reliable data. For example, an mHealth app may be relying on poorly drafted 'standard contractual clauses' (SCCs) for international data transfers (making it non-compliant), yet still be able to provide scientifically sound data to the users. If the scope of the exercise were an overall quality assessment of the mHealth app itself, then indeed, privacy could be a component to consider. However, as this exercise is about the data an mHealth app produces, it is irrelevant.

Data protection law is a complex field of European law. Including all aspects of data privacy into the assessment questionnaire will be entirely unfeasible as it would require hundreds of very specific and complex questions. It is unreasonable to expect a 'health care professional' to be able to make a proper assessment of whether an app adheres fully to privacy law or not. A simple question saying 'does the mHealth app adhere to privacy law' is a difficult question for any one person (who is not a privacy expert) to answer as an app may adhere to many principles, but perhaps not all. While it is the duty of each and every individual to notify the competent data protection authority if they feel that privacy rules are not being adhered to, only national data protection authorities are able to make a certified statement saying whether an app does or does not adhere to privacy law.

Transparency: This criterion overlaps heavily with the above issue of privacy. Many of the 'criteria' that Working Group members believe should fall into this field is actually covered by data protection law. Indeed, data protection law requires the following to be provided to the data subject:

- The identity and contact details of the controller (i.e. the mHealth app creator/owner)

- The contact details of the apps DPO (when applicable)

- The purposes of the processing for which the personal data are intended as well as the legal basis for processing (including legitimate interest)

- The recipients or categories of recipients of the personal data (if any)

- The transfer of personal data and which instrument is used (when applicable)

- The period for which the data will be stored

- The existence of the right to request from the controller access to and rectification or erasure of personal data, or to object to the processing of the data at any point as well as the right to portability of the data

- The existence of the right to withdraw consent at anytime

- The right to lodge a complaint with a DPA

- Whether the provision of personal data is a statutory or contractual requirement

- The existence of automated decision making, including profiling

Safety: "The criterion needed to assess safety relies heavily once again on a legal assessment. mHealth apps need to adhere to general safety requirements covered by the General Product Safety Directive. The issue of safety is an important one…. (but)…..issues related to safety of the app has little to do with assessing the reliability/validity of the data that an mHealth app produces. On the other hand, we consider that any safety element directly linked to the data should be covered under the data reliability and validity dimensions."

Interoperability: The ability of two or more eHealth systems to use and exchange both computer interpretable data and human understandable information and knowledge.–"It could be an area of work for the future (…) including semantic interoperability, could be an area of work for the future. *Note*: the aspect of data integrity during the exchange of data between several systems is assumed to be covered under validity.'

Technical stability: The technical stability of an app has little to do with whether the app produces reliable and valid data. This issue is more of a market based issued. If an app is technically unstable, it will likely not survive on the market place as it will lead to a lack of usage, poor reviews, and no downloads. If this exercise were to focus on the overall quality assessment of an app, then its underlying technical stability could be considered as a potential criterion.

Effectiveness: It refers to the health outcomes and as such should be excluded from the guidelines. 'It is likely that by the time an apps 'effectiveness' is fully assessed a new version of the app has been created making the assessment outdated'.

Usability/desirability: It is difficult to provide proper criteria here. As one person might view an app as 'desirable' because it is 'free', while another person might 'desire' the app because its colour scheme includes their favourite colour. 'The issue of usability and desirability is a market related issue. What makes an app usable or desirable is a complex discipline that cannot be easily summarised in a series of questions that will always have a level of ambiguity and subjectivity. Most app aggregators supply usability guidelines for app developers, and there is a risk that these guidelines could contradict those well established and accepted practices.'

Scalability: 'is purely a market related issue. The objective of this activity should not be to prescribe the platform of an app and how it should be dimensioned. If a mHealth app is unable to scale to a certain size, that is the choice of the app developer. Either the app will be forced to change so that it can compete on the market place or it will die.

There was no support for inclusion of any criterion other than reliability and validity (of data).

2.4 Public Authorities

Privacy: public healthcare centres cannot recommend using an app to their patients, if it requires a lot of personal information not necessary to use the app. A risk is perceived that

developers of the app could potentially make use of this such information in a questionable way

ACSA (Andalusia): A very significant number of health apps do not comply with privacy regulation, as documented in numerous research studies and scientific publications. Relying on a self-assessed and voluntary code of conduct does not seem to be enough.

(There is) no need to duplicate efforts in redefining different criteria from the Code of conduct, but the(re is a) need to externally assess them in some way.

INFARMED: We should bear in mind that we have two kinds of software products on the market (one regulated as a medical device and other that is not a device), the same technology that can be used/can compete on healthcare environment/settings, on the care pathway of the patient/interacting with EHR. Thus, a basic common level of quality criteria should be guaranteed from those that are not regulated as medical devices. These comments from Infarmed also apply to the criteria for Transparency, Safety, Reliability and Usability / desirability.

Transparency: It is essential that the user understands what information (the patient) is giving to the app. *ACSA (Andalusia)*: there are several aspects to be covered ... absence of conflict of interests, authorship, funding sources, etc. Apps developed by pharma, which provide clinical information, might be confused with drug publicity.

Safety: We can't put at risk, even to the smallest extent, the safety of the user of the app, even more if this risk arises from a misinterpretation of the data entered

ACSA (Andalusia): A risk assessment should be carried by the app developer, in order to respect the main principle: "First, do not harm"

Reliability: Stability should be also addressed within this criterion.

Validity: It is important that professionals contribute their knowledge in the clinical validation phase development of the app, as developers may not know well enough the disease treated by the app.

Interoperability: Data generated by a user must be able to be transmitted to their doctor, family or others who use different information systems.

Technical stability: *NHS (Glasgow)*: It's difficult for us to see a clear distinction between reliability and technical stability.

Effectiveness: A study of effectiveness requires a lot of time and during the intervening period, the version of the app evaluated can become obsolete and could be a new improved version.

AQUAS: We understand certification and assessment of the effectiveness as a longitudinal process, where first a certification is needed. The certification should guarantee performance against several important criteria before testing effectiveness. In an effectiveness study, the app should not be a prototype. During assessment of effectiveness, at a minimum, it should be demonstrated that the mHealth solution does not cause any harm to the users.

NHS (Glasgow): We read this as efficacy. This is important and linked with validity. Where validity provides a measure of the success of the system under test conditions, efficacy would provide a measure of desired or intended result in every-day use and particular environments.

Efficiency: *AQUAS*: For instance, the capacity to reach a vast part of the population with minimum resources invested, cost and time savings for the health system, to improve general health of the population.

NHS (Glasgow): We're not sure about this. How would a claim be separated from fact or experience?

Usability / Desirability: NHS (Glasgow): Usability is very important for us. Desirability is subjective, and is not so important for us.

Scalability: The most crucial aspect of scalability is being able to easily capture the results in the patient's EHR.

2.5 Payers and Social Health Insurance

Privacy: Privacy is essential, in order to comply with patient rights. "Privacy" consists of two areas:
1. Data protection: The rules of the EU Data Protection Regulation should be applied to mHealth apps. It is crucial, that data that are health related, thought they are collected by apps whose purpose is not medical, are also treated with the strict requirements the regulation defines.
2. Data safety: the apps should be safe against data breach etc. according to up-to-date standards

Transparency: The users must know what he gets. There are several topics regarding transparency:
1. Who is behind the app, which means, who distributes it, who financed it and who developed it. There must be contact data of the body which is responsible for the content of the app.

2. Where do the contents and the scientific concept of the app come from? Where can the scientific documentation of the app and its algorithms be found?

3. Provision of comprehensive information what data collected by using the app and how the operator of the app platform will use this data, e.g. for marketing or research purposes. All information on 1, 2 and 3. should be in clear language, well-structured and short, so that it is well understandable and accessible for the average user.

Safety: Basic principle: not to harm the user. The use of an app must be safe for the user in all its functions, i.e. without causing harm, irrespective of the data collecting function. E.g. the app is useless if it collects precise blood glucose data for the EHR (electronic health record) but makes mistakes when calculating the insulin dose for the user.

Reliability: A basic requirement - without reliability using an app makes no sense. Reliability may probably be defined as outlined by the ISO/IEC 25010 "A set of attributes that bear on

the capability of software to maintain its level of performance under stated conditions for a stated period of time." This item overlaps with technical stability

Validity: If data is meant to be used by medical professionals it must be valid. The same applies to the user. The data produced by the app must be either valid compared to a gold standard or it must make transparent under which circumstances it will produce less valid data and the extent of the error in this case.

Interoperability: If data is meant to be used by medical professionals the app must be interoperable. The app must be capable to exchange data with the most common electronic health records.

Technical stability: Overlaps with "reliability".

Effectiveness: The primary purpose of an app is to fulfil its function for the user. Collecting valid data to be used by third parties is a second order in most cases.

Accessibility: It makes no sense, if an app produces fine data but the users have difficulties using the app. All users of the target group of an app must be able to use it. In case there are limitations, the app should make them transparent, e.g. persons with low vision.

2.6 Research and Academia

Privacy: This is an important criterion for researchers since they are ultimately responsible for the secure management of all data collected in research projects.

A core challenge for mHealth device makers is to design solutions that allow target group users to safely disclose sensitive information to trusted external parties. Wearable devices seem to give the owner a sense of control that unfortunately, does not extend to the way in which it deals with the personal data it collects. The current crop of devices is black boxed solutions preventing personal data management.

Transparency: Same as above. For researchers to ensure research data is handled appropriately, they need a clear understanding of the data protection mechanisms that are in place. Also, for researchers to be able to stand over the data they are collecting, it is important to clearly understand the processes and sources that underlie the analytics that take place within mHealth solutions.

Current applications provide little clarity to end-users on how they deal with the personal data they collect. Although most apps include lengthy Terms of Services, virtually no information is provided as to the way data is collected in apps, the workings of the analytics and algorithms or the reasoning and steps that translate data to information and advice for end-users.

Transparency-by-design enables users and external observers to keep a close watch on the origin, quality, neutrality, value of data, the measures that have been taken into account to protect users' privacy and the processes and sources that underlie the analytics that take place within mHealth solutions. Trust can thus be improved by incorporating these approaches, increasing the likelihood of sustained and high quality user participation.

Safety: It is the researchers' responsibility to protect the well-being of research participants. In order to incorporate mHealth apps into research projects, it is vital to ensure the app won't cause harm to participants, be it of physical, mental, social or financial nature.

Reliability and Validity: The present generation of apps collects data by means of sensors, questionnaires and self-reporting. Reliability has consequences for the way in which the app interprets the user's behaviour and for the advice that is based on the data and the algorithms. Many apps give the impression to be reliable and valid, but the question is whether they could live up to their claims. Is the data being interpreted correctly, is the data based on scientific evidence or standards, which models or theories are being used and what is the consensus in science about these theories / models? Lack of compliance with these criteria leads to risk of wrong measurements and advice.

Interoperability: Recent years have seen a tremendous growth in so-called quantified self-management applications fuelled by the wide introduction of smart watches, home automation systems and other Internet of Things devices. The apps operate on a platform provided by the app-provider with different business models like free and premium (paying) users. They interface with sensors gathering all types of sensitive data, which is then - invisibly - analyse and aggregated and served back by the app provider. End-users typically have no control over these data streams data and unknowingly agree with the conditions of the app-provider with respect to data ownership.

Technical stability: Technically unstable systems have high risk of attrition

Effectiveness: A major barrier on the uptake of mHealth solutions within EU's healthcare systems, especially for healthcare providers, is the paucity of reliable evidence on the clinical benefits and cost-effectiveness of mHealth solutions. And even for fields of application where a number of studies exist, they are often not comparable (different study designs, inclusion criteria etc.) Studies on mHealth fail to provide persuasive results.

Usability / Desirability: Usability/ desirability plays an important role in adherence to the app, which, in turn, is important for the effectiveness of the app. Poor usability / desirability will impact participant engagement in research.

Scalability: For research projects with a larger study population, scalability of the apps (and supporting infrastructure where applicable) is an important factor to keep in mind. Especially for apps specifically developed within research projects/studies (e.g. by the research team), this is not always adequately honoured.