HAUTE AUTORITÉ DE SANTÉ

MEASURING
& IMPROVING QUALITY

STANDARD

Assessment of apps in the mobile health (mHealth) sector -Overview and quality criteria of medical content for referencing digital services in the digital health space and the professional service package

Validated by the Board on 24 June 2021

Description of the publication

Title	**Assessment of apps in the mobile health (mHealth) sector-Overview and quality criteria of medical content for referencing digital services in the digital health space and the professional service package**
Work method	Simple consensus method
Purpose(s)	Prepare an overview of assessment in the mHealth sector and set out the medical content quality criteria with a view to referencing apps in the digital healthcare space
Targets concerned	Health professionals and associations of healthcare professionals and patients involved in the mobile app assessment sector. Ministerial delegation for digital healthcare (DNS)
Requester	Ministerial delegation for digital healthcare (DNS)
Sponsor(s)	*Haute Autorité de santé* (French National Authority for Health) (HAS)
Project management	Coordination: Mr Pierre Trudelle, HAS Digital health commission (MNS) project manager (interim Head of Department: Marc Fumey) Secretarial duties: Ms Chantal Hottin and Ms Sophie Duthu
Literature search	Ms Marie Georget and Ms Laurence Frigère Head of Documentation - Monitoring Department: Ms Frédérique Pages
Authors	Mr Pierre Trudelle, project manager
Declaration of interests	The members of the review group have not declared any conflicts of interest.
Validation	Version dated 24 June 2021
Other formats	Criteria summary file

This document and its bibliographic reference are available to download at www.has-sante.fr

Haute Autorité de santé – Communication and Information Department
5 avenue du Stade de France – 93218 SAINT-DENIS LA PLAINE CEDEX. Tel.: +33 (0)1 55 93 70 00

Contents

Introduction 7

1. Context 9

1.1. Definition and development pathway of mHealth 9

1.2. Challenges to be addressed 9

1.3. Classifications 10

1.4. Recent development 11

1.5. The different domains of mHealth assessment 12

1.5.1. Convergence of different sectors and different actors 12

1.5.2. Developer assessment systems 12

1.5.3. Healthcare professional assessment systems 13

1.5.4. User assessment systems 14

1.5.5. Assessor assessment systems 14

1.5.6. Assessment challenges 15

1.6. Quantitative evolution of scientific literature 15

1.7. Levels of evidence and assessment of clinical utility 16

2. Potential risks of basic functions in the mHealth context 18

2.1. Role of risk in mHealth assessment 18

2.2. Most common basic functions 19

2.2.1. Inform 20

2.2.2. Instruct 20

2.2.3. Record 20

2.2.4. Calculate/analyse 21

2.2.5. Remind/alert 21

2.2.6. Display 21

2.2.7. Guide 22

2.2.8. Communicate 22

2.2.9. Variation of risk according to basic functions of mobile apps 22

3. International mHealth assessment methods 24

3.1. Scientific publication standard for mHealth-related publications 24

3.1.1. JMIR app submission form 24

3.1.2. CONSORT-EHEALTH and mERA standard 24

3.1.3. Literature review methodology for specific apps 24

3.2. The different mHealth assessment approaches/tools 25

3.3. Scores, checklists and other rating scales 27

3.3.1. Medical content quality score 27

3.3.1.1. Silberg scale or score 27
3.3.1.2. Brief DISCERN 28
3.3.1.3. Summary 28
3.3.2. MARS general score 29
3.3.2.1. MARS-G in Germany 29
3.3.2.2. u-MARS 30
3.3.3. Other general app quality scores 30
3.3.4. Specific app quality scores 30
3.3.4.1. Specific scores 31
3.3.4.2. Specific scores developed by learned societies 32
3.3.5. Specific app quality guidelines 32
3.3.6. Summary 32
3.4. User reviews, expert reviews 34
3.4.1. MyHealthApps-MHA (Great Britain) 34
3.4.2. GGD Appstore (Netherlands) 35
3.4.3. Health Navigator (New Zealand) 35
3.4.4. Onemind (USA) 36
3.4.5. Observatory: App sanitarie (Italy) 36
3.4.6. Summary 37
3.5. Labels/certifications with a list of criteria 37
3.5.1. ORCHA Review (Great Britain) 38
3.5.2. TICSS (Catalonia region-Spain) 39
3.5.3. Dekra certification (France) 39
3.5.4. DiaDigital seal (Germany) 39
3.5.5. HON code (Switzerland) 40
3.5.6. Summary 40
3.6. Editorialised repository consisting of a database providing a specific annotation of each app 41
3.6.1. Assessment guidelines for European countries 41
3.6.2. eHealth Suisse assessment domains 42
3.6.3. NHS Health Apps Library (Great Britain) 43
3.6.4. AppSaludable (Andalusia region, Spain) 45
3.6.5. Portugal 45
3.6.6. Summary 46
3.7. Editorialised repository consisting of a database providing a specific annotation of each app in a specialised sector 46
3.7.1. *Our Mobile Health* (OMH) (Great Britain) 46
3.7.2. Mental health with AppChecker (Denmark) 47
3.7.3. AppCheck (Germany) 47

3.7.4. Summary 48
3.8. Editorialised repository consisting of a database providing a specific selection of apps defined as medical devices 48
3.8.1. mHealthBelgium (Belgium) 48
3.8.2. DiGA (Germany) 48
3.8.3. Digi-HTA (Finland) 49
3.8.4. Summary 50
4. Medical content quality criteria 52
4.1. Different types of health content 52
4.2. Use of medical content quality criteria 52
4.3. Initial content design 53
4.3.1. Information management 53
4.3.1.1. Information service organisation 53
4.3.1.2. Expertise of authors of content in app 53
4.3.1.3. Declarations of interest 53
4.3.1.4. Quote of key sources and bibliographic references 53
4.3.1.5. Update of key sources and bibliographic references 54
4.3.1.6. Level of evidence 54
4.4. Generated content and standardisation 55
4.4.1. Relevance, trustworthiness of information and real-life justification 55
4.4.1.1. Measurement quality in context of use 55
4.4.1.2. User support 55
4.5. Interpreted content 56
4.5.1. Interpretation typology 56
4.5.1.1. Human interpretation of health content 56
4.5.1.2. Automated interpretation of health content 56
4.6. Displayed content 57
4.6.1. User understanding and involvement 57
4.6.1.1. Involvement of users (patients, professionals, specific parties) 57
4.6.1.2. Description of end-use 57
4.6.1.3. Contraindications, risks, limitations of use 57
4.6.1.4. Understanding of health content 58
4.6.2. Content display performance 58
4.6.2.1. User help/instructions 58
4.6.2.2. Text and image readability and navigation 58
4.6.2.3. Error prevention and understanding of information 58
Table of appendices 59

Bibliographic references 77

Participants 82

Abbreviations and acronyms 83

Introduction

This document has been drafted in response to the referral by the Ministerial delegation for digital healthcare (DNS) dated July 2020, requesting the HAS to draft an overview of the medical content quality criteria used in the mobile health (*mHealth*) sector.

This list of criteria, combined with other criteria (Appendix 1), will be used to contribute to the referencing of digital services to be listed in the digital health space (ENS)[1] and, by extension, for the digital service package intended for health professionals (BSP).

Update of the mobile health guidelines published in 2016

This contribution by the HAS is based on the guidelines published in 2016 (1), entitled: “Good practice guidelines on health apps and smart devices (*Mobile Health* or *mHealth*)”. At the time, the guidelines were aimed at manufacturers and assessors (assessment bodies, consumer associations or medical learned societies) who could and still can apply these guidelines to conduct their own assessments. They only applied to apps and smart devices with no declared medical end-use (often referred to as the “grey area” of apps or smart devices having a potential health effect without being a medical device). Note that to help clarify the boundary between medical devices and other products, the French National Agency for Medicines and Health Products Safety (ANSM) has published an update[2] on the categorisation of apps potentially falling within the remit of CE marking or not and being classified as a medical device. In this document, the term “app” covers all apps whether they are a medical device (MD) or not.

Document structure

This document contains two parts.

A **first part** (sections 1 to 3) giving an overview of mHealth assessment through some concrete examples. This part is not intended to be exhaustive, but is aimed at documenting the different assessment strategies and methods used internationally, and at helping develop information in this area. The articles selected describe assessment systems with a description focused on criteria. Very often, the type of app (MD or non-MD) that can be assessed with these tools is not specified. This also helps gain an understanding of the different assessment requirement levels according to the app type and its context of use.

The **second part** (section 4) provides a list of medical content quality criteria based on the 2016 HAS publication (1). The members of the working and review group who contributed to the 2016 document were contacted to review this version, along with additional peer reviewers.

[1] https://esante.gouv.fr/virage-numerique/feuille-de-route

[2] https://ansm.sante.fr/documents/reference/exemples-de-logiciels-et-applications-mobiles-illustrant-le-positionnement-reglementaire

Positioning

These guidelines are no substitute for medical device, privacy, and consumer protection legislation or regulations. The application of the good practices set out in these guidelines is intended without prejudice to the regulations in force.

The medical content assessment criteria proposed in this document apply to all types of apps. In the case of MD-classified medical apps, some proposed criteria may already be covered by the CE-MD mark process.

The HAS would like to point out that the CE mark is mandatory for digital services classified as MDs. This requirement could be taken into account in the referencing procedure by identifying the criteria corresponding to CE mark requirements and considering that, for the referencing of digital services classified as MDs, these criteria are deemed to have been fulfilled.

The HAS good practice guidelines are not an assessment tool for inclusion in lists qualifying for reimbursement.

1. Context

1.1. Definition and development pathway of mHealth

Mobile health is a field of eHealth (2). According to WHO, it covers "medical and public health practice supported by mobile devices, such as mobile phones, patient monitoring devices, personal digital assistants (PDAs), and other wireless devices[3]".

This field is currently dominated by "Mobile health and wellness apps" which are supposed to improve our quality of life, wellness, and health through, among other things, feedback on our **behaviours** (3, 4) and **targeted information**.

These mobile apps can connect with **medical devices** (e.g.: blood pressure monitor), **biosensors** (e.g.: smart wristband) or **data analysis systems** (e.g.: data management program/algorithm). The data collected could potentially be entered directly in the electronic patient record.

The term "*Mobile Medical Application*" (MMA) is frequently used in the literature to specify apps which are medical devices. This usage is not yet routine practice.

Each app generally targets a very specific and individualised objective, but the sector as a whole claims to enhance the quality of care in general, increase efficiency, and improve clinical research.

Furthermore, mHealth aims to increase user autonomy and responsibility (encompassed within the internationally used term *empowerment*). This empowerment does not replace healthcare professionals, but helps supplement treatment. This should help optimise referrals to healthcare professionals.

As regards chronic diseases, it is worth noting that the integration of mHealth is sometimes referred to as "*minimally disruptive medicine* – MDM", as it aims to reduce the "burden" of care for the patient (MDM-*based mHealth intervention*) (5, 6) *via* the use of specific apps or smart objects.

1.2. Challenges to be addressed

According to ORCHA (*Organisation for the Review of Care and Health Applications*), a company specialised in reviewing apps (with its system known as ORCHA *Baseline Review* – OBR), there are three main challenges to mHealth[4]:

- **Awareness** – Digital health apps and solutions as a whole are not yet part of the day-to-day management of health and care related conditions.
- **Accessibility** – Identifying apps in current general app stores is very difficult. Frequent search terms such as diabetes or depression refer to a number of apps on most sites, but they are generally merely a tiny proportion of the total number of apps actually available for these conditions.
- **Trust** – The lack of a suitable quality indicator inhibits the embracing of digital health by users, patients and healthcare professionals. The most common concerns affecting trust relate to: personal data confidentiality, safety, and clinical efficacy of the product. While many health apps cover a highly complex range of regulations and standards, there is no single consolidated service for viewing the overall conformity and compliance of an app and a clear indication for use.

[3] https://apps.who.int/iris/bitstream/handle/10665/44607/9789241564250_eng.pdf?sequence=1&isAllowed=y (page 6)
[4] https://www.orcha.co.uk/the-challenge/ (viewed on 05/11/2020)

ORCHA's comments are quite a good summary of the process of changing and embracing new technologies in medical settings and contexts.

Smart objects are specific entities and personal data protection and communication security aspects systematically apply (7).

1.3. Classifications

Very specifically and solely focusing on the medical care context, Kumar (8), in 2013, presented the results of a workshop in the United States on developing mHealth and how to approach it, to classify, and assess Apps. It proposes four sections integrated as part of a continuum: measurement, diagnostic, treatment/prevention, and global. In this classification, the apps may or may not be medical devices.

More recently, in 2017, the IQVIA Institute (9) classified the range of different solutions not categorised as medical devices offered to patients based on the following classification:

1. **Wellness & prevention** category:
- exercise and fitness;
- diet and nutrition;
- lifestyle and stress;
- stress management;
- sleep/insomnia;
- smoking cessation;
- alcohol moderation.

2. **Symptom onset and seeking care** category (patient experience tools):
- general healthcare information;
- symptom checking;
- finding a clinician;
- managing clinical and financial information;
- social media.

3. **Diagnosis** category:
- the healthcare professional may recommend an app to support the treatment programme;
- connected sensors for remote monitoring;
- apps for any use case across the patient journey.

4. **Condition monitoring** category (condition education and management):
- self-monitoring;
- remote patient monitoring;
- app-enabled rehabilitation programme.

5. **Treatment** category (prescription filling and compliance):
- Prescription discounts;
- prescription filling;

- medication management and adherence.

This structure is not exhaustive, but supplements the classifications listed in the 2016 HAS guidelines (1). It shows the “concrete” extent of the field of mHealth as at 2017.

Moreover, in 2021[5], the HAS published a **functional classification** according to the end-use of digital solutions used in medical or paramedical contexts. Four levels are proposed, ranging from system services for patients to autonomous decision management after data analysis.

In conclusion

One of the major challenges consists of finding a way to ensure a sufficient level of quality relative to the risks involved in the use of these mobile apps or smart objects.

1.4. Recent development

According to the data of the IQVIA Institute (9) from 2017, the proportion of apps targeting “wellness” dropped from 73 to 60% between 2015 and 2017. This means that 40% of apps targeted user health improvements in 2017, and that the wellness/health ratio is shifting towards the “health” sector.

According to IQVIA, the therapy areas in which digital solutions are available are ranked as follows:

- mental health and behavioural disorders (28%);
- diabetes (16%);
- heart/circulatory system (11%);
- nervous system (7%);
- musculoskeletal system and connective tissue (7%);
- cancer (5%);
- respiratory system (5%);
- digestive system (4%);
- eyes and ears (4%);
- pain (4%);
- endocrine, nutritional and metabolic diseases (3%);
- skin and subcutaneous tissue (3%).

The organisation ORCHA[6] quotes additional figures:

- over **327,000 apps are listed**;
- 43 apps represent 83% of downloads;
- 65% have not been updated in over eighteen months;
- over 80% of apps have fewer than 5000 downloads;
- under 7% of diabetics use an app for diabetes management, and under 2% of patients suffering from chronic obstructive pulmonary disease (COPD);
- over 240 conditions are covered;
- only 15% of apps reviewed by ORCHA meet the minimum assessment criteria.

[5] https://www.has-sante.fr/jcms/p_3238360/fr/classification-fonctionnelle-selon-leur-finalite-d-usage-des-solutions-numeriques-utilisees-dans-le-cadre-de-soins-medicaux-ou-paramedicaux

[6] https://www.orcha.co.uk/the-challenge/

Furthermore, Fagherazzi (10), in 2019, proposed the term “digitosome” which encompasses the set of digital data generated by a patient to be included in medical follow-up.

Finally, the mHealth sector claims that it can help deliver personalised medicine by offering the right treatment (through continuous indicator or targeted data analysis), for the right patient (through similar profile and pattern analysis), at the right time (through remote and real-time analysis), while helping develop clinical research (through data analysis and aggregation).

Regular use of mHealth apps among given cohorts (diabetes prevention, diabetes monitoring, asthma, cardiac and pulmonary rehabilitation) would, according to the IQVIA Institute (9), help save $7 billion a year (primarily by reducing emergency hospital admissions).

The PwC (PricewaterhouseCoopers) report (11) commissioned by the GSMA (*Global System for Mobile Communications*) in 2013 projected €99 billion in savings for the European Union in 2017 (above all through better prevention).

Finally, some targeted studies have tried to model the economic benefits of app use (e.g.: improving quality of life in the mental health sector (12)).

1.5. The different domains of mHealth assessment

The growth rate of the mHealth sector has led to a shift from “conventional” assessment systems, which has been described as “disruptive” by a European Union working group[7]. Over the last five years, we have seen an extremely wide range of assessment approaches in respect of these digital solutions (label, certification, online repository, scores, etc.). This can be explained by various reasons, in particular, by assessment frameworks specific to different actors in the sector (13-16).

1.5.1. Convergence of different sectors and different actors

The growth in mHealth is catalysed by the convergence of technological development in the telecommunications sector (Bluetooth, Wi-Fi, 5G), information technology and electronics (smartphone, electronic component miniaturisation, data processing and storage) supported by, as a minimum, four categories of actors:

- engineers;
- IT specialists;
- individuals and patients;
- healthcare professionals.

Each category of actors uses assessment systems linked with their own sectors. These assessment systems target specific objectives (technical trustworthiness, signal processing, public health, clinical benefit, cybersecurity, etc.) and are also restricted to the scope under assessment. Some examples are provided below to illustrate this.

1.5.2. Developer assessment systems

Engineers and IT specialists conventionally apply “process” type assessment systems:

- economic market access systems (e.g.: CE mark for medical devices);
- production process certification (e.g.: ISO standard);
- design standard use (e.g.: HL7 standard);

[7] http://ec.europa.eu/newsroom/document.cfm?doc_id=45251

- connectivity technology standards (e.g.: *Continua Design Guidelines* – CDG).

As regards standards, several will probably play an important role in the years to come. The list includes, among others:

- ISO 14971: application of risk management to medical devices;
- ISO 27001: information technology. Security techniques. Information security management systems. Requirements;
- IEC 62304: medical device software. Software life cycle processes (contains ISO 13485).

In 2021, the standard ISO-82304-2 (health software – Part 2: Health and wellness apps – Quality and reliability)[8] is undergoing validation on a European level, and supplements the standard ISO-82304-1 (health software – Part 1: General requirements for product safety).

The standard ISO-82304-2 is partly based on the British standard PAS-277 and specifically relates to mobile apps with 81 items. **This standard could help harmonise app quality requirements** internationally and reduce the proliferation of different assessment systems in different countries.

Moreover, regarding the CE marking process of medical devices, medical apps are subject to the new European Regulation 2017/745[9] on medical devices, which came into force on 26 May 2021, superseding the European Directives previously in force.

1.5.3. Healthcare professional assessment systems

Healthcare professionals use assessment systems based on epidemiological methodologies (e.g.: randomised controlled trial for treatment efficacy) or on qualitative approaches used in public health (17).

In 2013, Kumar (8) presented the different assessment methods that could be used according to the type of app proposed.

In 2019, the World Health Organization (WHO) (18) issued guideline recommendations for steering and assessing digital solutions and helping harmonise practices with summary tables compiling the different methodological approaches based on the objectives of the clinical assessment of these apps.

Pham (19), in 2016, demonstrated that the randomised controlled trial (RCT) is the most extensively used approach in mHealth assessment. She also showed that a fraction of the products used by consumers used this approach. Alternative methods should be proposed to adapt over time, and to address “sociotechnical systems”. This was confirmed by Byambasuren (20), in 2018, with an overview of systematic reviews on the efficacy of mHealth products. She highlighted the restricted scope of randomised controlled trials.

Mohr (21), in 2013, proposed developing assessment methods to assess the impact of behavioural intervention technologies. He proposed a framework known as CEEBIT (*Continuous Evaluation of Evolving Behavioral Intervention Technologies*) to assess the impact on users’ behaviour over time.

Grundy (22), in 2016, described the multidimensional aspects to be taken into account in this sector, which involves a “multiservice” assessment and the creation of innovative approaches to adapt to the sector.

Ologeanu-Taddei (23), in 2020, summarised the complexity of the methodologies that can be used in the mHealth sector. She mentions the scope of risk assessment which is still underestimated.

[8] https://www.iso.org/news/isofocus_141-6.html
[9] https://eur-lex.europa.eu/legal-content/FR/TXT/PDF/?uri=CELEX:32017R0745

As such, focusing the assessment solely on the clinical benefit can give rise to risks in other aspects. This is what occurred with the *NHS Apps Library* which listed apps online, two-thirds of which transmitted unencrypted personal data over the Internet (13, 24, 25). The promotion by the NHS of apps with these security vulnerabilities resulted in the library being closed and a thorough review of the assessment system for referencing.

Wierda (26), in 2020, published a personal data risk analysis according to the different collection and transmission phases in respect of these data. This privacy and security aspect is regularly cited as a barrier to trust in the use of apps and smart objects (7, 27).

1.5.4. User assessment systems

Users or patients who are not experts in the matter use external resources to assess mobile apps or smart objects (28). Most often, users rely on the opinions of professionals or those close to them, or user reviews on the manufacturer's website.

Note that the latter ratings may be subject to caution. Pustozerov (29), in 2016, observed an average rating of 4 out of 5, regardless of the type of app rated (standard deviation of 1.65). Furthermore, 90.7% of ratings are distributed among just 4.1% of available apps which limits the benefit of this type of rating to the most popular apps. Finally, it is advised to be wary of fake reviews which are sometimes paid for by manufacturers and which should be distinguished from certified reviews. The French Directorate General for Competition and the Repression of Fraud (DGCCRF) has issued a guide sheet on the topic[10] with an ISO standard.

Singh (30) demonstrated the poor correlation between user ratings and clinical utility (Spearman rho of 0.21; p = 0.02) or the System Usability Scale (Spearman rho of 0.11; p = 0.2).

Some documents providing guidance on how to choose apps or the precautions to be taken have been published for the general public[11,12,13,14].

1.5.5. Assessor assessment systems

There are many actors involved in assessment. The first private-sector actors sell proprietary assessment systems (generally in label format), the assessment criteria of which are not in the public domain (referred to as assessment "black box"), and thus represent their trade secret[15].

For international institutional actors, involved in setting up app assessment systems, user data security is a major priority.

Learned societies and patient associations try to recommend specific solutions for their members and, as a general rule, in specialist fields with the best benefit/risk ratio for specific patient profiles.

These approaches are currently faced with two barriers, either the assessment system is overly stringent and the list is not useful and poorly populated, or the assessment system is too general and is not specific enough to be reliable and trustworthy for users.

[10] https://www.economie.gouv.fr/dgccrf/Publications/Vie-pratique/Fiches-pratiques/faux-avis-consommateurs-sur-internet
[11] https://www.economie.gouv.fr/dgccrf/Publications/Vie-pratique/Fiches-pratiques/objets-connectes
[12] https://gskpro.com/content/dam/global/hcpportal/fr_FR/AiresThrapeutiques/pneumonew/pdf/info-patients.pdf
[13] https://www.mentalhealthcommission.ca/sites/default/files/2018-01/eMH_app_fr.pdf
[14] https://www.healthnavigator.org.nz/apps/h/how-to-choose-a-health-app-a-guide-for-clinicians/
[15] https://dumas.ccsd.cnrs.fr/dumas-02971517/document (pages 83-84)

1.5.6. Assessment challenges

Finally, the viewpoints of actors, assessors or users cover a large number of assessment domains and assessment procedures, which complexifies assessment systems which attempt to globalise all of the approaches in this sector.

Assessment challenges generally target the following domains:

- technical challenge (e.g.: trustworthy software and/or hardware design);
- cybersecurity challenge (e.g.: cyber risk assessment process [16], security vulnerability management);
- privacy challenge (e.g.: compliance with EU Regulation 2016/679 of the European Parliament and of the Council of 27 April 2016 on data protection [General Data Protection Regulation – GDPR] and with the European "*Code of Conduct on privacy for mHealth apps*"[17,18]);
- medical quality challenge (e.g.: trustworthiness of medical content);
- user interface performance challenge[19] (e.g.: literacy, navigation, ergonomics, clinical scenario tests);
- economic challenge (e.g.: economic model viability).

1.6. Quantitative evolution of scientific literature

In parallel, there is exponential growth in research studying the utility of mHealth. Ali (31) studied the evolution of publications between 1995 and 2015. He showed a growth in articles, and categorised the sectors studied into five domains:

- health promotion and prevention;
- diagnosis;
- treatment;
- monitoring;
- healthcare service support.

Articles on **patient monitoring** apps and **diagnostic** apps showed the strongest growth since 2013. These domains are assessed with clinical studies compiled in systematic reviews. Note that Fiordelli (32), in 2013, had already predicted the evolution of research towards greater measurement of the health impact of apps.

Furthermore, our literature search on guidelines and systematic reviews showed the ramping up of clinical publications on this sector. This **ramping up has gathered pace in the last five years** showing that the sector is looking to develop clinical assessment by seeking levels of evidence. However, the quality level remains mixed (33, 34).

[16] https://a51.nl/sites/default/files/pdf/Pacemaker%20Ecosystem%20Evaluation.pdf
[17] https://ec.europa.eu/digital-single-market/en/news/code-conduct-privacy-mhealth-apps-has-been-finalised
[18] http://ec.europa.eu/newsroom/dae/document.cfm?action=display&doc_id=16125
[19] https://www.oracle.com/technetwork/topics/ux/applications/user-profile-template-1884987.pdf

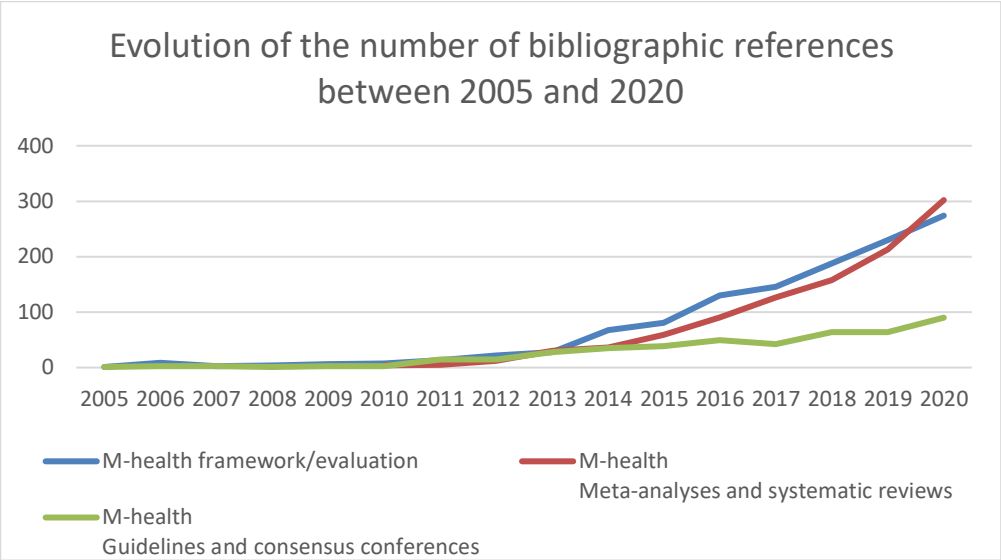

Figure 1. Evolution of the number of publications in the mHealth sector from 2005 to 2020

1.7. Levels of evidence and assessment of clinical utility

The IQVIA Institute (9) summarised the trends between 2007 and 2017 in publication type (234 randomised controlled trials and 20 meta-analyses) and categorised the sectors in which the body of evidence was growing. As such, according to the IQVIA Institute, the areas of diabetes, depression and anxiety are sectors in which apps should be listed in clinical practice guidelines.

IQVIA lists **24 categories in which the quality of publications make apps strong candidates for adoption**:

- weight management/healthy eating;
- asthma;
- infectious & parasitic disease;
- chronic obstructive pulmonary disease (COPD);
- hearing loss & tinnitus;
- CHF;
- stroke;
- Alzheimer's disease;
- medication management;
- alcohol & substance abuse;
- sleep/insomnia;
- cancer;
- smoking cessation;
- Parkinson's disease;
- hypertension;
- cardiac rehab;
- stress management;
- alcohol moderation;
- PTSD;
- genitourinary conditions;

- diabetes prevention;
- arthritis;
- kidney disease;
- pulmonary rehab.

On the other hand, according to IQVIA in 2017, other sectors were disappointing or require further studies:

- exercise;
- pain management;
- dermatological conditions;
- bipolar/schizophrenia;
- multiple sclerosis;
- autism.

IQVIA presents a timeline of the type of publications: from the observational study to the first clinical trial(s) to meta-analyses with an improving quality level over time. This is a natural progression, as there is a minimum duration between the set-up of a study and its publication, and this sector has only seen around ten years of actual growth.

As regards the use of apps in clinical research, specific websites are available, such as the *Clinical Trials Transformation Initiative* – CTTI[20]. It proposes guidelines for the use of apps in the context of clinical trials[21,22]. At the end of 2020, the CTTI database had a list of 438 clinical trials[23].

A query run on the United States clinical trials registry[24] in 2021 identifies over one thousand intervention-based clinical trials with the keyword “mHealth”.

Finally, international conferences solely focusing on mHealth and its impact in clinical trials are being organised[25].

In conclusion

The interest generated by apps has given rise to a demand for assessment with a two-pronged objective of improving the quality of apps and increasing users’ trust.

The mHealth assessment process is multidimensional. It is especially based on standardisation of production, personal data privacy and protection, IT and telecommunication security, management and prevention of the risks generated by app use, the public health impact, user uptake, and the economic model for app maintenance and development.

Hence, after an initial flurry in the creation and production of apps covering wellness, fitness or health, a phase in which scientific publications and clinical studies have ramped up has been in progress since 2013, with a qualitative improvement in the publications. However, the multitude of assessment systems complexifies the implementation of an overall system.

[20] https://www.ctti-clinicaltrials.org/news/ctti-update-advancing-use-mobile-health-technology-transform-clinical-trials
[21] https://www.ctti-clinicaltrials.org/sites/www.ctti-clinicaltrials.org/files/ctti_recommendations_-_mct_engaging_patients_and_sites_final.pdf
[22] https://www.ctti-clinicaltrials.org/sites/www.ctti-clinicaltrials.org/files/mobile-technologies-executive-summary.pdf
[23] https://feasibility-studies.ctti-clinicaltrials.org
[24] https://clinicaltrials.gov/ct2/results?term=mhealth&age_v=&gndr=&type=Intr&rslt=&Search=Apply
[25] https://impacct-mhealth.com/about/what-to-expect-from-a-digital-event/

2. Potential risks of basic functions in the mHealth context

The use of mobile apps or smart objects gives rise to different types of risk according to the use and claimed functionalities, the user profile, and the context of use.

2.1. Role of risk in mHealth assessment

The concept of risk and risk assessment can be approached in matrix form, used to link up two or more dimensions after identifying a threat or hazard. For example, with a “frequency/severity” analysis following a frequency analysis of incidents/accidents and effects.

Quantitative (probabilistic) or qualitative analyses also exist (35).

Lewis (36), in 2014, identified the different levels of risk in mHealth. In his view, the risk is above all linked with **app complexity**, as this increases risk. Therefore, he defined two dimensions in his risk matrix:

- from the simplest app (low risk) to the most complex (high risk);
- from a low chance of harm (low risk) to a high chance of harm (high risk).

The types of risks are identified and ranked in this paper with scenarios liable to cause these risks.

According to Bradway (37), in 2017, the risk dimensions proposed in the mHealth context are also dependent on the **user’s profile**. He cites:

- intervention-specific risk (reference & guide, communicate & coordinate, diagnose & treatment, monitor & alert);
- person-specific risk (healthy living, chronic and moderated conditions, severe conditions, frail and at-risk).

This approach targets the individual use of apps and smart objects in the human dimension. Apps are developed to meet users’ specific needs. This may generate risks in the **context of use** (38).

In Germany, the IGES report (39), in 2016, proposes a risk matrix for healthcare products that are potentially medical devices with two dimensions:

- from the user’s perspective;
- based on the level of functional autonomy of the product (from general information to the replacement of care providers by algorithms).

In Australia, the TGA (*Therapeutic Goods Administration*) (40), in 2020, conducted a literature review on the risk of apps. They were identified in general and also for specific health issues (symptom checkers, diabetes management, melanoma/skin analysis, asthma, cardiovascular measurements, medicine dose). This report underlines the **poor risk assessment level and the potential impact on decision-making**.

For its part, the HAS (1), in 2016, defined a two-dimensional risk matrix for apps which are not medical devices:

- from the perspective of the primary use of the product (from information to targeted data analysis);
- from the primary user’s perspective (from the general public to healthcare professionals).

This choice is based on the risk of misuse or inappropriate use of the product, especially for apps which are not medical devices.

The risk is high if healthcare professionals make decisions based on incorrect information, as this would impact large numbers of patients.

For **medical apps which are medical devices, the risk class is assessed and the appropriate requirements need to be met**[26].

In conclusion

Assessment should be modulated by the risk associated with app use (see the benefit/risk ratio of app use).

This risk level can be grasped by taking the following three dimensions into consideration:

- the **proposed end-use of the app** which determines the (intrinsic) hazard in the event of harm associated with normal functioning or malfunctions;
- the **context of use of the app** which determines the ability to act upon the risk and may, depending on the circumstances, help prevent it from occurring (detection) or lessen its effects (mitigation, recovery);
- the **target population** using the app which determines the accepted effects and may be a potential risk factor (misuse, incorrect interpretation or comprehension).

2.2. Most common basic functions

Mobile apps can have one or more end-uses claimed by the manufacturer. They range from providing general information to recommendations specifically prepared based on user-provided data. There are more or less specific risks associated with the different types of app “end-use”. In addition to this risk, there is a risk of personal data exposure to unauthorised third parties or due to cyberattacks as in the case of the NHS store in 2015 (41).

Kearney (42) proposed a system for modelling end-uses according to healthcare sectors and user type. Two lines of approach are defined:

- an **end-use** line of approach including: information/communication, assessment, intervention, monitoring, coordination/teamwork;
- a **health** line of approach including: wellness, prevention, diagnosis, treatment, follow-up (monitoring).

The intersection of these two lines of approach gives a view of the purpose of the app for the general public, patients, or healthcare professionals.

Schematically, the **basic functions** observed most often[27] in the mHealth sector are as follows:

- inform;
- instruct;
- record;
- calculate/analyse;
- remind/alert;

[26] https://ansm.sante.fr/page/mise-sur-le-marche-des-dispositifs-medicaux-et-des-dispositifs-medicaux-de-diagnostic-in-vitro

[27] https://www.iqvia.com/-/media/iqvia/pdfs/institute-reports/patient-adoption-of-mhealth.pdf (pages 7-8)

- display;
- guide;
- communicate.

These basic functions are used individually or together to achieve the aims of the app. Some produce, disseminate or use personal data with a risk of exposure to unauthorised third parties (30). Assessing this risk is complex (43).

The order of presentation below arbitrarily increments increasingly advanced functions and refers to all user types (users, patients, carers, healthcare professionals).

2.2.1. Inform

This consists of **general and identical health information provision and access, regardless of the user**. The information relates to healthy living and health promotion, health education, accessibility to information in formats tailored to the user, patients, or healthcare professionals. This information is sometimes "pushed/delivered" automatically to the user, when they have selected specific topics on which they wish to receive information. They may also wish to be alerted or "notified" of new publications or findings on selected topics or gradually addressed topics (44).

2.2.2. Instruct

This consists of **educating the patient or providing more specific instructions for the user or healthcare professional**. The use of questionnaires, games (conceptualised by the term *gamification* (45)), real-life scenarios, exchanges via interactions with a virtual environment or a group of persons of reference is the means used most often to adapt the information to be provided according to the user's knowledge or query. For healthcare professionals, this may involve guidance on performing a specific procedure.

2.2.3. Record

This consists of **tracking or capturing health data** (e.g.: monitoring photos of the progression of a mole into melanoma, anxiety levels in the course of the day, and any collection of "digital biomarkers" to help predict the potential onset of specific health issues or conditions, etc.).

It may also consist of **measures entered by the user** (e.g.: urination schedule for better management of urinary intervals in the case of urinary disorders, completing regular questionnaires, reaching planned targets, etc.). The best-known areas are related to exercise, weight, and diets, or parameters associated with sleep quality or mood.

In the context of medical devices, the parameters measured help real-time or deferred remote tracking of the user's health. The best-known measures are cardiovascular parameters (e.g.: blood pressure, heart rate, oxygen saturation, peak flow meter, etc.), body temperature during postoperative or infection monitoring or monitoring of chronic conditions (e.g.: stable blood glucose for patients with diabetes, etc.).

Measures can sometimes be **obtained in the form of questionnaires**. For example, PROMs (*patient-reported outcomes measures*) to assess treatment outcomes, or PREMs (*patient-reported experience measures*) to assess the experience of and satisfaction with care provided.

2.2.4. Calculate/analyse

This consists of **making calculations based on patient data** (e.g.: mathematical operation to calculate doses according to weight, scores, etc.) or data analyses (with human interpretation or interpretation *via* algorithms). The quality of the calculation is linked with the data collection quality (46).

As regards algorithms, European-wide guidelines were published in 2019 (47). In 2021, the European Commission is working on setting an artificial intelligence (AI) regulation[28]. A proposed risk level-based assessment approach is used for different AI systems. For the high-risk level, AI systems could be subject to the following strict requirements:

- adequate risk assessment and mitigation systems;
- high quality of the datasets feeding the system to minimise risks and discriminatory outcomes;
- logging of activity to ensure traceability of results;
- detailed documentation providing all information necessary on the system and its purpose for authorities to assess its compliance;
- clear and adequate information to the user; appropriate human oversight measures to minimise risk; high level of robustness, security and accuracy.

Autonomous AI systems could be registered in a European Union database. To be introduced onto the market, a declaration of conformity could be required and the AI system would bear the CE mark. This specific field will undoubtedly need to be assessed specifically.

Numerous publications are available setting out the risks associated with calculation and data analysis (24, 46, 48-50).

2.2.5. Remind/alert

This consists of **issuing alerts or reminders** intended for the user or for a clinician contact. Initially provided *via* SMS messages aimed at the user for medication doses, vaccination or screening reminders, user reminders and alerts have evolved into specially designed apps using different formats of reminder (notification, display and specific message, etc.) and tracking for exercise, mental health, smoking cessation encouragement, geolocation for allergy risks (pollen, vaccination, etc.), etc.

For healthcare professionals, alerts are configured for patient follow-up, prevention and monitoring of specific risks (e.g.: remote medical monitoring).

In the context of medical devices, the monitoring parameters measured are used to alert the patient and clinician contact remotely (falling within the scope of remote patient monitoring [RPM]). Alert levels are configured to position the alert thresholds and types and the parties concerned by an incoming alert (51).

2.2.6. Display

This consists of **displaying, in one or more display formats, the results of the data collected or processed by the app**. The different display modes are aimed at enhancing the information sent to the user to increase the effectiveness of the message sent. These data may have been recorded by the user, captured by sensors integrated in the smartphone, collected by a smart object/medical device or produced by a computing algorithm.

[28] https://eur-lex.europa.eu/procedure/FI/2021_106

2.2.7. Guide

This consists of **providing guidance** based on the information collected from the user, or potentially recommending a specific action, a consultation with a particular clinician or a treatment when the system targets the healthcare system user.

When the tools are developed for healthcare professionals, they are generally designed as a decision-making aid.

2.2.8. Communicate

This consists of **developing communication between healthcare professionals and healthcare system users**. The tools may be synchronous (e.g.: real-time communication during a telecare consultation) or asynchronous (communication via separate messages).

2.2.9. Variation of risk according to basic functions of mobile apps

If the basic functions of an app are ineffective, they will generate risks. These risks are characterised by their likelihood and by their effects (according to the frequency/severity model).

As regards medical or health content, the most common effects are an **unsuitable time to treatment** due to false reassurance or poor information, or, on the other hand, overmedication or unnecessary user anxiety. The level of quality of the health information and data used impacts care optimisation (defined as providing the right care, at the right time, for the right patient).

Table 1 contains the basic functions of apps in the health sector with examples of types of risks and their effects on health content. Further examples of associated risks are reported to provide context and illustrate the potential effects more broadly.

Table 1: Basic functions of an app and examples of risks

Functions	Examples of risks for medical content	Other examples of risks
Inform	Risks of misinformation. Depends on information quality level: source, authors' credentials and conflicts of interest, level of evidence, currency.	Risks of incorrect interpretations. Depends on level of understanding and relevance of selection by the user of information channels used.
Instruct	Risk of bad practice. Depends on information quality level, but also on teaching or educational methods used in the app.	Risk of selection bias. Depends on performance level of user needs assessment and of relevance of proposed responses.
Record	Risks of recording poor-quality data whether from the user, or *via* an external device. Depends on trustworthiness of data entered, collected, recorded and archived.	Risks associated with access or modifications of health information and data collected from user (cybersecurity, personal data protection, technical glitches).
Calculate/Analyse	Risks of generating calculation or analytical errors. Depends on trustworthiness of data recorded to perform calculations, but also trustworthiness of calculation methods used and interpretation performance level.	If interpretation is performed *via* an algorithm, the risks are associated with maintaining algorithm performance in different clinical scenarios and use settings. The assessment of systems referred to as a whole as "artificial intelligence (AI)" is to be adapted according to the type of AI used.
Remind/Alert	Risks of not generating suitable alerts (absence, delay, too many alerts, etc.). Depends on settings entered to trigger	Risks associated with collecting information and defining alert thresholds or alert type. This information can be entered or amended by a third

	alerts and interpretation of notifications by users.	party (security risk) or by the user (risk of misuse) and be unsuitable for the user's circumstances.
Display	Risks of generating a deviation between actual data and those displayed. Depends on extent of deviation between raw data collected and how they were processed for display (smoothing, granularity, calibration in relation to standards/cohorts, image quality, conversion to movement pattern, etc.). This deviation may no longer reflect the reality of the measure studied.	Risks associated with understanding. Depends on level of understanding of data display (by user, by carer or by professional) and resulting effects on decision-making. Risk associated with the health data "sharing function" to boost or encourage the user (personal data protection).
Guide	Risks associated with relevance of proposed guidance. The proposed guidance may impact decision-making and give rise to unsuitable care by increasing human costs (e.g.: hospital admission, inappropriate test, stress) or economic costs (e.g.: unnecessary expenditure).	If interpretation is performed *via* an algorithm, the risks are associated with the quality of the rationale proposed by the system (which information are the suggestions, guidance, decisions based on? Guaranteed by whom?.
Communicate	Risks associated with relevance of responses given.	Risks associated with delays in responses for organisational or technical expertise reasons.

In conclusion

An app generally provides one or more basic functions (inform, calculate, guide, communicate, etc.). These functions can be encountered at different stages of the user or patient pathway (wellness and prevention, follow-up and monitoring, treatment, etc.).

The trustworthiness of an app's basic functions determines to an extent the quality of the health information and data collected. The user should be informed of the risks associated with using the app (description of potential effects) *via* suitable indicators or targeted communication means.

3. International mHealth assessment methods

Depending on the countries, the mHealth assessment system(s) are based on one or more of the approaches listed below.

3.1. Scientific publication standard for mHealth-related publications

The innovative development of mHealth apps has impacted the format of research publications in the field. The *Journal of Medical Internet Research* (JMIR) has proposed an app assessment grid similar to the submission of an article for a journal (with peer review). Publication standards or standards for publishing summary reviews on apps have been defined. The format of publications has been structured in this way over the last ten years.

Note that an international consensus on use of the term “app” was published by Lewis (52) in 2014 for app-related scientific articles published in English. The acronym MMA (*mobile medical application*) is frequently used to refer to apps which are medical devices.

3.1.1. JMIR app submission form

When drawing up the 2016 guidelines (1), the HAS referred to the research by Riezebos (53). He had conducted an overview of the literature to compile the quality criteria to be used to assess mHealth apps.

His research resulted in the creation of a **grid of criteria for the original app submission for review** JMIR (2013). This form, which is still available online[29], allows apps to be submitted for peer review based on a number of aspects (level of evidence, safety, etc.) based on the same procedure as for the review of an article.

3.1.2. CONSORT-EHEALTH and mERA standard

Eysenbach (54), in 2011, developed a checklist to be used when submitting a trial for publication relating to an app or other similar types of interventions in the eHealth sector. CONSORT-EHEALTH is referenced in the EQUATOR (*Enhancing the QUAlity and Transparency Of health Research*)[30] network which compiles authors’ guidelines for researchers and scientific publishers. It is currently the only standard available for app-related trial publications (as at 21/11/2019).

Note that Agarwal (55), in 2016, developed a checklist of just 16 items to simplify the submission process for some apps (particularly those from developing countries). The project was supported by WHO. The checklist is known as mERA (*mHealth Evidence Reporting and Assessment*).

3.1.3. Literature review methodology for specific apps

Several hundred controlled trials have been published in the mHealth sector. Boudreaux (56), in 2014, proposed a specific literature search strategy to assess the body of literature on the utility of apps in specific health domains, but also by exploring various other information sources.

[29] tinyurl.com/appsform

[30] https://www.equator-network.org/reporting-guidelines/consort-ehealth-improving-and-standardizing-evaluation-reports-of-web-based-and-mobile-health-interventions/

He proposed seven steps:

1. **Review the scientific literature.** Search the scientific literature for papers reviewing apps in a content domain or strong clinical trials.
2. **Search app clearinghouse websites.** Clearinghouses that review apps can help with identifying strengths and weaknesses.
3. **Search app stores.** App stores are challenging to navigate. It is important to fine-tune and filter app searches with the most relevant and targeted key words.
4. **Review app descriptions, user ratings, and reviews**. Publicised ratings and user reviews can offer evidence of app usability, functionality, and efficacy, which can help to narrow the pool of candidate apps.
5. **Conduct a social media query within professional and, if available, patient networks**. Social networks may reveal new app trends, likeability by certain user groups, and other substantive data.
6. **Pilot test the app.** Apps may be piloted by the healthcare provider or a designee, including examinations of functionality, accuracy of content, and usability.
7. **Elicit feedback from patients.** Patients may be able to provide valuable insights after they have used the app a provider recommends.

This pragmatic strategy would provide an overview of the use of apps for specific health domains or specific issues identified, for example, by learned societies.

3.2. The different mHealth assessment approaches/tools

A large number of authors and organisations have conducted literature reviews to combine the (heterogeneous) assessment approaches proposed internationally in the mHealth sector. Some authors cite the 2016 HAS guidelines in their overview (57-64).

Among these publications, *via* the federal body eHealth Suisse [31], Switzerland published an international literature review of mHealth assessment methods (15) in 2019.

This body coordinates eHealth governance in the country, and has produced various documents[32] (guidelines, quality criteria, linking of apps, technical specifications and standards, etc.).

It proposes to categorise the results of this literature review according to three levels, with the following structure, adopting and adapting the CHARISMHA classification (65).

➔ Assessments linked with general tools/instruments

These are defined by eHealth Suisse (15) as “various tools or various measures, for example a guide with a checklist and/or editorialised mobile app repositories”.

Two categories are included in this conceptual assessment category:

- **regulations** including certificates of conformity and accreditations (e.g.: CE mark for medical devices);
- **codes** including a list of criteria or rules in respect of good conduct or quality proposed on a voluntary use basis.

[31] https://www.e-health-suisse.ch/fr/page-daccueil.html
[32] https://www.e-health-suisse.ch/fr/mise-en-oeuvre-communautes/activites-ehealth/mhealth.html

➔ Third-party assessment

This is defined by eHealth Suisse (15) as an external assessment conducted by "experts in a specific medical specialty or technical staff, and not by manufacturers or users themselves. However, experts can also gain access to self-reported manufacturer data".

Six categories of tools are included in this conceptual assessment category:

- **codes with a view to assigning a label** including the same types of codes as above, but assessed by third parties. Meeting the criteria potentially results in a label being assigned (the checking procedures can vary in degrees of stringency);
- **quality labels** generally assigned by businesses or organisations primarily from the private sector (generally concerning apps not covered by the CE mark or not subject to regulations);
- **editorialised repositories** or **assessment platform** generally organised in database format, listing a predefined selection of apps. According to the repository, the apps displayed are endorsed with one or more assessments conducted by medical experts, technical experts, or users;
- **expert reviews** and **personal reviews** left on online stores in the form of comments or star ratings. Reviews are sometimes published in report form;
- **methodical tests** are targeted assessments with known or partially undisclosed criteria conducted by businesses or organisations providing these services;
- **scientific studies** are generally systematic reviews or meta-analyses of studies published on a specific type of apps with similar endpoints.

➔ Individual assessment by interested parties

These are defined by eHealth Suisse (15) as an independent way of assessing, to support users or managers when selecting or deciding for or against the use of an "mHealth" solution.

- **Standardised product information descriptions** present the relevant app information homogeneously and transparently. This information is published in suitable locations by the manufacturer (product information through online store, product website, etc.).
- **Practical assessment instructions** provide a checklist in the form of a list of criteria to aid users in assessing the app. This assessment format is used to calculate a score or a ranking based on the calculated results.

This classification will be used below to structure the overview of the various international assessment approaches identified in the literature review.

In conclusion

The technical challenge involved in assessment is found on two fronts:

- Which assessment **areas** are examined effectively and pragmatically?
- Which **assessment** and **organisational strategies** need to be put in place?

There are two possible strategies: inclusive (less strict) or selective (stricter) tool models.

3.3. Scores, checklists and other rating scales

Scores and checklists have been devised to facilitate app assessment. Aimed at users or professionals, these tools can be used to assess apps quickly, independently and easily, by assigning a score or identifying the expected components of the app with a checklist.

Azad-Khaneghah (66), in 2020, conducted a literature review on rating scales in the mHealth sector. He listed 23 scales assessing app usability and 25 scales assessing app quality (87 publications selected and analysed in total). Azad-Khaneghah underlines the ambiguity of the term quality which is interpreted differently according to the authors or assessed in different or restricted domains.

3.3.1. Medical content quality score

In 2007, the HAS (67), published a literature review on assessment methods in respect of eHealth sites and the quality of information circulated on the Internet. This review demonstrated, at the time, the variety of scales assessing website content. The same observation can be made in relation to the assessment of mHealth app quality.

Some scales cited in the 2007 HAS report (67), and used for websites, have been adapted for assessing apps. Van Singer (68), in 2015, thus cited the Abbott, Brief DISCERN (six-item version of DISCERN rated from 1 to 5), Health On the Net (HON) scales, or the Silberg score.

Butcher (69), in 2015, created a 100-point score with three domains (content, sources and levels of evidence) to assess apps producing medical content. This score is especially suitable for apps listing scientific publications.

3.3.1.1. Silberg scale or score

A large proportion of the medical content assessment scales available in eHealth refer to the editorial published in 1997 by Silberg (70), Editorial Director at JAMA, on assuring the quality of medical information on the Internet. In this editorial, Silberg singles out four factors for rating information quality:

- **authorship**: cite authors and contributors. Cite their contributions and their credentials;
- **attribution**: cite references and sources of information for all content. This should be listed clearly along with any relevant copyright information;
- **disclosure**: the owner(s) should be clearly identified in full; the same applies for sponsors, any patronage, advertising, advertorials, commercial funding arrangements or similar support, and any potential conflicts of interest. This also includes arrangements linking and referring to other sites for a financial reward. This type of standard should also be extended to discussion forums;
- **currency**: content publication and update dates.

This editorial has since been used and the Silberg scale constructed with dedicated and adapted scores.

For example, Zhang (71), in 2017, assessed the content of 14 apps focused on postnatal depression. The average Silberg score was 3 (+/- 1.52) out of 9 points, which enabled Zhang to highlight the lack of disclosure of information sources. The details of the Silberg score used in this study are listed below.

- **authorship** (3 points): whether authors are identified, whether affiliations are identified, whether credentials are identified;
- **attribution of information sources** (2 points): whether sources are given, whether references are given or hyperlinked in text;
- **disclosure** (2 points): whether app ownership is disclosed, and whether sponsorship is disclosed;

- **currency** (2 points): whether app has been modified in the last month, whether app has included a last modification date.

According to the users of the Silberg scale, adaptations to the breakdown of points are proposed (68, 72).

This scale essentially measures the degree of disclosure of the information and potential conflicts of interest. It is not suitable for rating the relevance of selected and circulated information.

3.3.1.2. Brief DISCERN

Brief DISCERN (73) is the short version of DISCERN[33] which includes 16 items. A 5-star version is also available (67).

Brief DISCERN (68) examines content in slightly more depth (compared to the Silberg score) with 6 items (rated from 1 to 5):

- is it clear what sources of information were used to compile the publication?
- is it clear when the information used or reported in the publication was produced?
- does it describe how each treatment works?
- does the publication describe the benefits of each treatment?
- does it describe the risks of each treatment?
- does it describe how the treatment choices affect overall quality of life?

The responses to these items are more descriptive, but do not allow a full assessment of the trustworthiness of the information selected.

3.3.1.3. Summary

Scores solely devoted to medical content quality generally focus on the publication of information based on the **four items of the Silberg scale** (authorship, source, ownership and conflict of interest, currency). These scores, which are too limited for most apps, are supplemented by scores covering several additional domains (Table 2).

Specific approaches assessing the content of a specific medical topic are more accurate for assessing content. They are constructed similarly to audit frameworks (74) with in some cases metrological assessments in respect of content validity and reproducibility or randomised controlled trials. They will be discussed later in this document.

[33] http://www.discern.org.uk/discern_instrument.php

Table 2: Examples of scores solely assessing medical content

Name	Country	Categories assessed	Score	Comments
Silberg score	USA	Authorship, affiliation, disclosure, currency.	Variable based on adaptations, but generally on a 9-point scale.	Different weighting depending on authors.
Brief DISCERN	DISCERN (Great Britain) Brief DISCERN (Switzerland)	Sources, information production date, how the treatment works, benefit of treatment, risk of treatment, choice of treatment and impact on quality of life.	6 items rated from 1 to 5.	Production of database containing apps assessed using this score.
Butcher score	Canada	Content, disclosure, level of evidence.	Percentage on 8 items.	Designed for apps identifying information resources.

Note that the NHS, in Great Britain, has set out a list of six principles and 16 items for assessing the information production process known as *Information Standard Principles*[34]. These principles are devised to demonstrate that the creator has implemented a clearly defined process for producing and maintaining high-quality, evidence-based, health and care-related information. This information should be available for the user.

These statement items enable a more in-depth analysis of the information production quality process. Appendix 2 is based on a non-validated French translation of this standard.

3.3.2. MARS general score

The MARS (*Mobile Application Rating Scale*) score is a composite score made up of several assessment domains. It was constructed using 349 items grouped into six categories following a literature review (75) in 2015. The score has been reduced to **23 items rated from 1 to 5 in four objective domains and one subjective domain**. The score has been adapted to several languages, including Italian (76), Spanish (77), German (78) and Arabic (79). A French translation was provided in the 2016 HAS guidelines; this French translation has not been validated to date (Appendix 3).

The metrological qualities of this score are good (80), and assessed for different types of apps (81). Dozens of publications have used (82) or assessed the use of this general score which is ahead of the many other scores or checklists, but remains limited in terms of the scope of its assessment and particularly the levels of evidence (83).

3.3.2.1. MARS-G in Germany

In Germany, the MARS score has been used to set up an app repository known as: "*Mobile Health App Database* (MHAD)[35]". This repository was created in 2018, and, as at the end of February 2021, included 1112 apps encompassing nine categories: mindfulness (192 apps), anxiety (104 apps), depression (39 apps), support for children and teens (13 apps), cancer (75 apps), PTSD (82 apps), pain (218 apps), support for senior citizens (77 apps), exercise (312 apps).

The assessment methodology was described by Stach (84), in 2020. Online stores (App Store and Google Play) are browsed using an automated process. The results are then filtered and divided into

[34] https://www.england.nhs.uk/tis/about/the-info-standard/#information-production
[35] http://www.mhad.science

suitable domains by reviewers. The MARS score has been modified with an additional section on "therapeutic gain" (gain for patients, gain for therapists, risks and side effects, ease of implementation into routine healthcare). This version was named MARS-G by the author and was validated in German by Messner (78), in 2020. Note that a specially developed video tutorial for users is available online[36].

Similarly, Australia[37] also uses the MARS score to provide a dedicated online repository[38] with over 300 apps claimed in 2020.

NB: an exhaustive search into similar initiatives in other countries has not been conducted.

3.3.2.2. u-MARS

Stonayov (85), in 2016, developed the user version of MARS. Made up of 20 items rated on a 5-point scale (with a quality section and a subjective section), the score contains an additional "perceived impact" section.

3.3.3. Other general app quality scores

Various initiatives have sought to broaden assessment beyond the remit of the MARS score. A non-exhaustive list given below shows the categories assessed while providing some details of the medical content-related section.

The *Royal College of Physicians* (RCP) Health Informatics Unit developed, in 2015 (86), an **18-item checklist covering three topics** to help doctors assess apps:

- who developed the app and what's inside it (9 questions)?
- how well does the app work (4 questions)?
- is there any evidence that the app does actually alleviate the problem (5 questions)?

ORCHA-24 (*Organisation for the Review of Care and Health Applications*) is a short version (87) of the ORCHA questionnaire. It was constructed based on regulatory sources, the PAS-277 standard, and expert feedback (according to the Delphi method). A more detailed description is provided in Appendix 4.

Jin (88), in 2015, developed a score based on a literature review and contributions from an expert panel. **Five categories were listed for 23 items rated from 0 to 3**:

- content (accuracy, comprehension);
- objectivity (authors, expertise);
- interface;
- term precision;
- technical security.

3.3.4. Specific app quality scores

The above scores are **"general" scores; they are constructed to be applicable to all types of apps**. Another approach can be used to assess medical content with a "specific" score. In this case, a scale is constructed to search for the key information of the topic in question.

[36] https://www.youtube.com/watch?v=5vwMiCWC0Sc
[37] https://www.pulseitmagazine.com.au/australian-ehealth/2660-bad-apps-and-where-to-find-them
[38] https://www.vichealth.vic.gov.au/media-and-resources/vichealth-apps/healthy-living-apps

3.3.4.1. Specific scores

As a general rule, authors identify the information to be disseminated in good practice guidelines, and apps are reviewed in the form of a systematic assessment (*systematic assessment of Apps*) to identify apps containing this information.

The criteria are based on good practice guidelines, evidence-based data, or psychological or behavioural factors (89).

These systematic assessments are generally conducted when the health issue is common and a large number of apps are available. There is often an educational aspect of the information or behavioural change.

Some published examples:

- **based on guidelines**: adherence to exercise guidelines (90, 91), adherence to a behavioural approach for tobacco cessation (92) or blood donation (93), adherence to eczema guidelines (94), user survey for pain management apps (95), expert assessment with 10 items rated from 1 to 5 enuresis-related apps (96);
- **based on behavioural change**: exercise promotion (97), 62 NICE criteria for behavioural change intervention apps (98).

In terms of validation of this type of scale, criteria are selected with support from specialists in the field (construct validity) and the reproducibility of the assessment is rated between reviewers (qualitative or quantitative statistical consistency).

Specific studies are sometimes used for monitoring guidelines *via* apps. Siebert (99), in 2020, developed an app (known as "*Guiding Pad App*") to apply paediatric cardiac resuscitation guidelines. The decision tree published by a learned society was broken down into several steps, and a randomised controlled trial assessed the efficacy of the use of this app.

Watson (100), in 2020, developed an app for tailored blood transfusion prescription. A review compared the improvement in the clinical decision with or without the app.

DiFilippo (101), in 2017, developed a quality score known as AQEL (*App Quality Evaluation*). After a literature review and expert opinions, he proposed a score providing an overall assessment (of the education and technical features of the app), and he added a supplementary section to assess specific topics. This score is designed for nutrition apps.

ACDC (*App Chronic Disease Checklist*) is a checklist implemented for the assessment of chronic disease-oriented apps by healthcare professionals (102). It is made up of four domains (engagement, functionality, ease of use, and information management) and 24 criteria. The objective is for professionals to inspect and test apps and check the presence of specific features.

The benefit of these type of guidelines is two-fold — firstly, they help target the health content that the app should contain, and secondly, they ensure that this information is provided, understood, and applied.

An example of the level of detail of the "clinical content" assessment for pain management apps is provided in Appendix 5 based on the article by Reynoldson (95) in 2014.

This example of expected clinical content criteria helps assess the sought requirement level for this type of app and the option of comparing apps with an identical reference.

3.3.4.2. Specific scores developed by learned societies

In 2014, the *American Society of Health-System Pharmacists* drafted an initial summary document (103), to provide checklists with a view to assessing medication-related apps (trustworthiness, dose calculated, etc.). Loy (104), in 2016, provided a more detailed description of the score targeting proposed features for medication-related apps (monitoring, drug interaction verification, dose calculation, medicinal product information, medicinal product recall and registration).

In the same vein, Camacho (105), in 2020, described a scale (TEACH-Apps) supported by the *American Psychiatric Association* made up of eight categories (privacy, medical evidence, price, ratings, attributes, features, onboarding, performance). The review targets mental health-related apps and is performed by a member of an expert committee who review apps in around thirty minutes.

Llorens-Vernet (63), in 2020, conducted a literature review of the various scores and compiled the criteria identified. A total of 503 criteria were identified and a selection of 36 criteria deemed to be important was retained for eight assessment categories.

3.3.5. Specific app quality guidelines

Besides scores which are often restricted to around twenty items, guidelines provide a more in-depth assessment and systematically cover several assessment domains.

Some examples are listed below, with more details provided when medical content is discussed.

Nouri (60), in 2018, published a literature review on criteria for assessing app quality. It lists, among other things, 15 studies including an assessment of medical content. A total of seven categories were drawn up with 37 subcategories.

Henson (106), in 2019, collected the criteria from 45 quality guidelines, and compiled the 604 items to obtain 357 which were studied in terms of their structural characteristics. 357 items were selected in total. A five-tier pyramid divided the assessment categories:

- level 1: background info;
- level 2: privacy and security;
- level 3: evidence based (first impressions after using, clinical validity, user feedback supporting);
- level 4: ease of use;
- level 5: data integration.

3.3.6. Summary

In sum, **scores provide a quick overview of apps with a general or more in-depth focus for specific scores**. Among these scales, the MARS (*Mobile Application Rating Scale*) score has been studied and used the most. Table 3 compiles the examples mentioned above, supplementing with other similar identified examples.

Table 3: Examples of composite scores assessing different mHealth domains

Name	Country	Domains assessed	Number of criteria	Comments
MARS (75)	Australia	Engagement, Functionality, Design, Information quality	23	Validation by various studies
MARS-G (78)	Germany	As for MARS, with added therapeutic gain domain	27	Database containing apps assessed using this score
No name (Jin)[39]	South Korea	Content (accuracy, comprehension), objectivity, interface, accuracy of terms, technology	23	
No name (Llorens-Vernet) (63)	Spain	Ease of use, privacy, security, relevance and adequacy, content disclosure, security, technical support and currency, technology	36 items	
QoE (Martinez)[40]	Spain	Content quality, security, ease of use, availability, performance, appearance, learning, accuracy	21 items	Perceived subjective assessment
No name (Robustillo)[41]	Spain	Design and relevance, information quality and security, service provision, confidentiality and privacy	40 items	Used to compare apps
AppScript (IQVIA)[42]	USA	Professional assessment, patient assessment, assessment according to six features, institutional support, development techniques, clinical assessment	ND	
APPLICATIONS (Chyjek)[43]	USA	Understanding of app, price, literature used, connectivity, advertising, research field, inter-device compatibility, other media components, ease of navigation, subjective presentation	16	Developed to compare apps
RCP (86)	Great Britain	Who developed the app and what's inside it? How well does the app work? Is there any evidence that the app does actually alleviate the problem?	18 items	Aids clinician in selecting apps
ORCHA-24 (87)	Great Britain	Data governance, clinical efficacy, user experience	24 items	Excerpt from a longer questionnaire
MedAd-AppQ (Ali 2018)[44]	Singapore	Content trustworthiness, feature utility, ease of use of feature	24	

ND: not defined

[39] https://www.liebertpub.com/doi/pdf/10.1089/tmj.2014.0151
[40] https://link.springer.com/article/10.1007/s10916-013-9976-x
[41] https://www.liebertpub.com/doi/pdf/10.1089/tmj.2013.0262
[42] https://www.appscript.net/score-details
[43] https://journals.lww.com/greenjournal/Fulltext/2015/06000/Rating_Pregnancy_Wheel_Applications_Using_the.29.aspx
[44] https://www.sciencedirect.com/science/article/abs/pii/S1551741117306654?via%3Dihub

Multipurpose guidelines provide over 300 criteria on average and are aimed at more specific assessment. Table 4 compiles the examples mentioned above, supplementing with other similar identified examples.

Table 4: Examples of multipurpose guidelines for assessing mHealth

Name	Country	Domains assessed	Number of criteria	Comments
Nouri (60)	France	Design, content (credibility, accuracy, information quality, information quantity), ease of use, functionality, ethics, security and privacy, user perception	ND	
Henson (106)	USA	Background info,	357	
Enlight[45]	USA	Usability, visual design, user engagement, content, therapeutic persuasiveness, therapeutic alliance, general subjective evaluation	476	Combination of five checklists: credibility, evidence-based programme, privacy explanation, basic security
ORCHA *Review* (OBR)[46]	Great Britain	Not detailed	260 to 350	

ND: not defined

3.4. User reviews, expert reviews

This was the first type of assessment encountered when the first apps emerged.

User surveys or feedback are conducted and studied. Van Haasteren (107) proposed the mHAT checklist (to survey users' opinions on app trust). This type of assessment is described below with a few representative examples.

3.4.1. MyHealthApps-MHA (Great Britain)

MyHealthApps-MHA is one of the largest databases of free and accessible apps (1160 apps listed categorised by medical specialty and by condition)[47]. Apps are submitted for assessment online[48] via a form available in nine languages. Apps are selected based on two main factors[49]:

- the health app has been nominated as a favourite by patient/disability/carer/family/consumer groups, or by empowered consumers (e.g. consumer advocates, active members/bloggers of moderated consumer health forums);
- the app developer is transparent about the nature of the app. App background checks are performed by PatientView and include various components (pricing, authenticity of the user assessment, contact details of the app designers and owners).

Each app has an overview and any recommendation by an association, learned body, or official structure.

[45] https://www.jmir.org/2017/3/e82/
[46] https://nhsprocurement.org.uk/wp-content/uploads/2020/11/Attachment-1-ORCHA-Baseline-Review-OBR-Process-.pdf
[47] http://myhealthapps.net
[48] http://myhealthapps.net/submit
[49] http://myhealthapps.net/methodology

3.4.2. GGD Appstore (Netherlands)

GGD Appstore is a database in which assessment is carried out by professionals and patients (229 apps listed, with the possibility of an app featuring in several categories). Apps are submitted for online assessment.

GGD Appstore offers six app categories (with the possibility of an app featuring in several categories)[50]:

- physical wellness;
- mental wellness;
- self-improvement;
- quality of life;
- engagement;
- everyday life.

Each app has a descriptive sheet and two tabs and a 0 to 5 star rating based on responses to a list of questions. A compilation document is available for download containing the description and responses to the questions asked. The general assessment methodology is published[51]:

- description;
- description of the app, its use and its prices;
- who is the app aimed at?
- what can the app be used for?
- what can the app do?
- where can it be downloaded?
- assessment (40 items);
- usability (15 items);
- trustworthiness (5 items);
- justification (5 items);
- confidentiality and security (15 items).

3.4.3. Health Navigator (New Zealand)

Health Navigator is **one of the first databases with structured criteria**, published online. This database is constructed based on an app search (219 apps listed in 68 different health issue or healthy living-related categories). The categories are determined by the Health Navigator editorial team[52], or *via* external clinician or consumer requests.

Health Navigator runs searches to locate apps. New apps are identified particularly by[53]:

- a literature review of research papers;
- published reviews of individual or categories of apps on other independent app review websites;
- searches on app stores;
- trending apps on social media and popular news;
- being alerted by app developers who can submit the app through an online form[54];

[50] https://www.ggdappstore.nl/Appstore/Homepage
[51] https://www.ggdappstore.nl/Appstore/Testmethode
[52] https://www.healthnavigator.org.nz/apps/a/app-library/
[53] https://www.healthnavigator.org.nz/apps/p/people-process/#App%20selection%20process
[54] https://docs.google.com/forms/d/e/1FAIpQLSdTd3amG2qAMj-JO5Rn5e2OtQzQnnQlFBKCI1wHO5JKgPJ9PQ/viewform

- through website users, including consumers using the apps.

Another literature search is then conducted to help identify useful features and criteria to assess apps within that category (e.g.: functionality features desirable in medicine reminder apps).

The app review process comprises four stages:

- internal review (functionality, quality of the information produced, target audience);
- New Zealand relevance review (unit of measurement, food product not available in New Zealand);
- clinical review (review by a healthcare professional working in the relevant area, clinical value, relevance, security. Rating of 1 to 5 assigned);
- user review (to find out whether the app does what a user expects it to, what the user likes or dislikes about the app).

Apps are excluded if they are deemed to be clinically unsafe or potentially harmful to users. Other reasons for excluding apps include incomplete content, functionality issues and security or privacy issues. The app name and the reason for exclusion are documented on the app category overview page.

3.4.4. Onemind (USA)

Onemind is a database specialised in a specific topic: mental health [55] (193 apps listed and 55 withdrawn, as no longer available to download). It has been set up by a professional association including a team of lead experts and a panel of specialist reviewers. Assessment is carried out in three areas rated out of a 5-point scale:

- credibility;
- user experience;
- transparency.

The name of the reviewers is cited for each reviewed application. Technical information and the app research status are also documented in the description of the reviewed app.

3.4.5. Observatory: App sanitarie (Italy)

This is a national repository set up in Italy in 2016 (App sanitarie[56]) to list apps covering twenty medical fields. A total of 640 apps, and 12 with the CE mark are listed and described.

The document is a compilation of apps. The information produced is mainly descriptive.

[55] https://onemindpsyberguide.org/apps/
[56] https://www.appsanitarie.it

3.4.6. Summary

Table 5 compiles the examples mentioned above, supplementing with other similar identified examples.

Table 5: Examples of websites featuring user and/or expert app reviews

Name	Country	Assessment system	Number of apps	Comments
MyHealthApps[57]	Great Britain	User assessment	1,160	
GGD Appstore[58]	Netherlands	Expert and user assessment	229 for the 6 categories	Some apps fall under several categories
Health Navigator[59]	New Zealand	Expert and user assessment	219	Some apps fall under several categories
Onemind[60]	USA	Expert assessment	193	55 withdrawn, as no longer available to download
Osservatorio APP sanitarie[61]	Italy	Expert assessment	640 (in 2016)	12 apps with CE mark in 2016
HealthOn[62]	Germany	Expert assessment	474	
Digimeda[63]	Germany	Expert assessment	Paid access	Assessment[64] on a scale of 1 to 7 (from controlled trial to no assessment). Website, chatbots assessed
Groupe Pasteur Mutualité[65]	France	Expert assessment	Claims over 800 reviewed apps[66]	Restricted access

The number of apps was counted at the end of February 2021

3.5. Labels/certifications with a list of criteria

Some public or private organisations have designed specific label or certification processes to help differentiate apps or highlight a specific assessment process (compliance with a list of criteria, technical tests, legal review, etc.). The label may be promoted by the manufacturer on various media.

The framework follows the conventional quality assurance and external quality control model. Service quality is generally defined as the ability to meet the user's needs safely.

[57] http://myhealthapps.net
[58] https://www.ggdappstore.nl
[59] https://www.healthnavigator.org.nz/apps/a/app-library/
[60] https://onemindpsyberguide.org/apps/
[61] https://www.appsanitarie.it
[62] https://www.healthon.de
[63] https://digimeda.de
[64] https://digimeda.de/ueber-digimeda#ranking
[65] https://www.gpm.fr/actualite.html?id=10093
[66] https://www.ticsante.com/story/2142/e-sante-une-application-recense-et-evalue-plus-de-800-dispositifs.html

3.5.1. ORCHA Review (Great Britain)

The private review organisation[67], ORCHA (*Organisation for the Review of Care and Health Applications*) was **one of the first bodies in Europe to implement a process combining a review engine and an expert review** (2015).

Designed by clinicians, ORCHA Review offers accreditation for digital health services worldwide. The basic review assesses 260 criteria (compliance in the areas of clinical safety, data confidentiality and user experience, regulations and international standards), and covers over 350 health issue categories.

Each review is tailored and customised according to the focus and functional capabilities of the app. The platform runs weekly automated reviews of all app updates.

Apps are ranked in five levels (from 0 to 4) based on their area of interest and their functional capabilities. The more an app focuses on "health" and is rich in features, the higher its ranking level and the greater the number of review domains activated for review. For example, very basic apps focusing on wellness (level 0 or 1) are not assessed based on their clinical performance, as they do not really provide a clinical solution unlike level 4 apps. The functional review also dynamically modifies the review fields monitored by reviewers during the actual review to ensure that the review is the best possible fit for the type of app.

An overall ORCHA score is constructed from the responses to each of the items in the assessment domains. Some items are positive and others negative. Thresholds of 65% (app with issues) and under 45% (potentially unusable or unsafe app).

If a new version is not issued within 18 months following the current version, the link to the app is marked "obsolete" and the ORCHA app score will start to decline at a rate of 5% per month.

The assessment framework also offers an in-depth assessment covering over 500 criteria in five domains (including financial and commercial stability, and country-specific custom adaptations).

ORCHA claims to be able to run hundreds of reviews each week, in over 180 condition and category domains. To date, ORCHA is reported to have reviewed over 6000 apps.

ORCHA reviews its criteria with international experts on a quarterly basis. It is currently running a review of the scale of the system with NeLL in preparation for 2030[68].

Critical app reviews are hosted on a platform allowing a smart search to identify apps against a specific range of criteria. The platform makes it possible to create bespoke app libraries according to clients' specific requests quickly and easily.

ORCHA conducts reviews for governmental organisations throughout Europe, the Middle East and Australasia.

In the United Kingdom, ORCHA conducts reviews for NHS Digital and the NHS in 50% of regions (as a national innovation acceleration programme).

[67] https://www.orcha.co.uk/our-solution/the-orcha-review/#0

[68] https://www.nell.eu/upload/images/news/ORCHA%20mHealth%202030.pdf

3.5.2. TICSS (Catalonia region-Spain)

TICSS is a foundation under the Catalan ministry of health which offers various services. In the mHealth sector, it offers an assessment service referred to as accreditation. The process involves professional associations and an expert committee[69].

It covers four blocks:

- **Usability and accessibility**. The application must have an intuitive interface, with a design suited to its intended function and it must ensure universal, inclusive access to people with functional diversity to maximise the benefits offered by the technology.
- **Technology**. The app functions efficiently and reliably from a technological point of view. The application must adapt to a minimum of functionality acceptable to the end user ensuring robustness and consistency.
- **Security**. Robust mechanisms are in place to preserve the privacy of the data generated by the users and the utmost confidentiality in the transmission of said information. It is necessary to ensure proper storage of information and establish mechanisms for encryption when registering passwords.
- **Functionality and content**. A committee of experts, made up of professionals from different professional associations in the sector such as professional associations, evaluates the quality of the content and the utility of the functions offered. Their review includes the usability and design of the application and whether it notifies the user of any software updates.

3.5.3. Dekra certification (France)

MedAppCare was the first body to be awarded an accredited certification level (accreditation No. 5-0598) by the French Accreditation Committee (Cofrac). It was taken over by Dekra[70] in 2020.

It assesses and certifies the quality of mobile apps, web platforms and connected health, wellness, autonomy loss and disability services.

Four domains are covered:

- data protection;
- digital security;
- relevance of content;
- ergonomics and use.

Once awarded, the certification is valid for three years.

3.5.4. DiaDigital seal (Germany)

The app manufacturer requests the seal[71,72] and conducts a self-assessment of its app (108). The Bochum Telematics and Telemedicine Centre (ZTG) runs a technical check and issues a report. The DiaDigital app testers conduct individual assessments. During a telephone conference, which all testers can attend, a check is conducted as to whether the app meets all the important criteria. The testers' feedback is summarised in a final report.

[69] https://ticsalutsocial.cat/en/serveis/mhealth-en/accreditation-service-and-ticss-guarantee-certification/
[70] https://www.dekra-certification.fr/certification-de-services/certification-applications-mobiles-et-sites-web-dekra-certification.html
[71] https://appcheck.de/bewertung-durch-diadigital-und-pneumodigital/
[72] https://www.diabetesde.org/diadigital

The app is published if a decision has been made on the choice of self-disclosure. The outcome of the technical review and the findings are available in the “Certified apps” section.

If an app does not meet the criteria, the manufacturer receives feedback on the scope for improvement and can then reapply.

3.5.5. HON code (Switzerland)

The HON Foundation which provided website reviews has developed an assessment process for apps: the HON code[73]. Described as a certification, this assessment system is somewhat similar to a label. It is made up of eight principles:

- authority (editorial team);
- complementarity (limitations of the app);
- confidentiality (legal requirement);
- validity (date of update);
- justifiability/objectivity (complete, objective references);
- user’s practice (usability and access to support);
- financial disclosure (funding sources and transparency of paid services);
- advertisement policy (identification and separation of advertising).

The assessment is conducted by medical or legal experts for data protection. HON code also provides automated vulnerability tests or confidentiality or encryption tests.

3.5.6. Summary

Table 6 compiles the examples mentioned above, supplementing with other similar identified examples.

Table 6: Examples of organisations providing labels, certification or accreditation

Name	Country	Assessment system	Number of apps
ORCHA	Great Britain	Label	Over 6000
TICSS	Spain (Catalonia)	Accreditation	4 (01/21)
Dekra (formerly MedAppCare)	France	Certification	ND
OMH (Our Mobile Health)[74]	Great Britain	Label	ND
DiaDigital	Germany	Label	ND
HON code	Switzerland	Label	ND
OMH (Our Mobile Health)	Great Britain	Label	ND

The number of apps was counted at the end of February 2021. ND: not defined

[73] https://www.hon.ch/fr/certification/app-certification.html
[74] https://www.ourmobilehealth.com

3.6. Editorialised repository consisting of a database providing a specific annotation of each app

Some public bodies provide app databases (repositories) adding a very in-depth and specific assessment in several domains.

3.6.1. Assessment guidelines for European countries

There are a variety of mobile app quality assessment domains. A **European network has been set up to identify European initiatives and convey the guidelines used**.

In 2016, at the same time as the HAS guidelines were being drafted (1), a working group organised by the European Commission was formed to draft European guidelines on assessment criteria for mobile apps and smart objects. This paper followed on from the Green Paper on mHealth published in 2014[75]. The objective was to determine the data that could be entered in the electronic patient record *via* these specific digital tools[76].

The working group was unable to reach a consensus, due to, simply put, two opposing leanings:

- one seeking a multi-domain assessment;
- one seeking an assessment solely on the trustworthiness of the data potentially entered in the electronic patient record.

Finally, the HAS guidelines were mentioned among the six European guideline projects (two from the UK, two from Spain, one from Germany)[77]. They are also mentioned by the European report on safety of non-embedded software published in 2019 (SMART 2016/071)[78].

Work on European mHealth guidelines has continued with support from the World Health Organization (WHO), the International Telecommunication Union (ITU), the Ministry of Health of the Region of Andalusia, and a grant from the European Commission with the "*European mHealth Innovation and Knowledge Hub*[79]". The HAS guidelines are mentioned in the report (currently being finalised at the time of writing of this document) as an assessment framework among 24 guidelines covering nine countries.

This report[80] provides a summarised timeline of the different assessment guidelines or systems used in Europe over the last twelve years.

> 2008: Continua Design Guidelines;
> 2012: AppCheck; MedAppCare; Andalusia region;
> 2013: My Health Apps; Our Mobile Health; NICE initiative;
> 2015: ORCHA; PAS 277; TicSalutSocial;
> 2016: DiaDigital; GGD Appstore; European guidelines; HAS;
> 2017: mindapps.dk; NHS digital
> 2018: APPKri; cMHAFF, PneumoDigital; MySNS Seleção; mHealthBelgium;
> 2019: AppQ; eHealth Suisse;
> 2020: BfArM; ISO initiative.

[75] https://eur-lex.europa.eu/legal-content/FR/TXT/HTML/?uri=CELEX:52014DC0219&from=EN
[76] https://ec.europa.eu/digital-single-market/en/news/report-working-group-mhealth-assessment-guidelines
[77] https://ec.europa.eu/digital-single-market/en/news/report-working-group-mhealth-assessment-guidelines
[78] https://op.europa.eu/fr/publication-detail/-/publication/aad6a287-5523-11e9-a8ed-01aa75ed71a1/language-en
[79] https://mhealth-hub.org/assessment-frameworks
[80] https://mhealth-hub.org/download/d2-1-knowledge-tool-1-health-apps-assessment-frameworks-pending-ec-approval (page 53)

mHealth Hub reviewed these 24 assessment guidelines and summarised the mutually consistent or inconsistent criteria and proposed 28 assessment recommendations for the domains analysed. The objective could be to enable the development of mutual recognition by comparing the criteria used by different countries.

The **12 assessment domains** analysed were as follows:

- privacy;
- transparency;
- safety;
- reliability;
- validity;
- interoperability;
- technical stability;
- effectiveness;
- accessibility;
- scalability;
- user experience and usability;
- security.

Note that the European Union is starting to provide specific checklist type tools for app developers (e.g.: FI-STAR[81]).

A further "ISO-oriented" route and assessment standardisation, mentioned above, is also being studied on a European level with several ISO standard, including one focusing specifically on apps (ISO-82304-2[82]) due to be published in September 2021.

European funding ("horizon Europe" programme) is available in 2021 for developing a European label[83].

3.6.2. eHealth Suisse assessment domains

eHealth Suisse, following a literature review, (15) propose **nine assessment domains** broken down into 18 characteristics and 25 requirements (109):

- transparency;
- fitness for purpose;
- risk proportionality;
- ethical acceptability;
- legal compliance;
- content validity;
- technical suitability;
- ease of use;
- resource efficiency.

[81] https://www.fi-star.eu/publications/checklist-for-app-developers.html

[82] https://www.iso.org/news/isofocus_141-6.html

[83] https://digitalhealtheurope.eu/resources/funding-opportunities/promoting-a-trusted-mhealth-label-in-europe-uptake-of-technical-specifications-for-quality-and-reliability-of-health-and-wellness-apps/

An example of the “content validity” domain is given below with the characteristics to be self-reported and the requirements listed:

- **content validity** (the content and the functions refer to valid sources. They are evidence-based in compliance with any applicable directives and reflect recent [scientific] knowledge):
  - **source validity**: it is requested to specify how content quality is guaranteed (e.g.: referral to experts in the field in question) and which valid sources are used (particularly accounting for new scientific knowledge, directives, studies, etc. specifying the level of evidence);
  - **content currency**: it is requested to specify how the app is regularly adapted visibly to new requirements in terms of content.

The benefit of this model is that it covers a large number of aspects and limits the amount of requirements. The system had not yet assessed an app in 2020.

3.6.3. NHS Health Apps Library (Great Britain)

Among the **first national repositories to be set up**, the *NHS Health Apps Library*[84] has undergone a number of upgrades.

The selection of the assessment domains and criteria to be used is based on the Digital Health & Care Institute report (58), in 2018, supported by the European programme, INTERREG. This report summarised the different health assessment approaches. It helped define the domains to be assessed:

- data protection;
- credible sources – evidence-based information;
- user experience and usability;
- functionality;
- authentication and security;
- impact – effectiveness;
- interoperability.

This report was used to set up the *Digital Assessment Questions* (DAQs) and *Digital Assessment Portal* (DAP) used by the NHS[85] to assess healthcare products until October 2020.

In February 2021, a new version of the assessment procedure was proposed. It is based on the *Digital Technical Assessment Criteria* (DTAC)[86] and comprises four technical domains and one domain covering key benchmarks.

The four technical domains:

- clinical safety;
- data protection;
- technical assurance;
- interoperability.

The fifth domain provides benchmarks in terms of app accessibility and usability.

Table 7 provides an example of the description of the first two criteria in clinical safety[87].

[84] https://www.nhs.uk/apps-library/
[85] https://digital.nhs.uk/services/nhs-apps-library
[86] https://www.nhsx.nhs.uk/key-tools-and-info/digital-technology-assessment-criteria-dtac/
[87] https://www.nhsx.nhs.uk/documents/60/DTAC_version_1.0_FINAL_updated_16.04.odt

Table 7: Examples of criteria described in NHS guidelines (Great Britain)

Code	Question	Options	Supporting information	Scoring criteria
C1.1	Have you undertaken Clinical Risk Management activities for this product which comply with DCB0129?	Yes/No	The DCB019[88] standard applies to organisations that are responsible for the development and maintenance of health IT systems. A health IT system is defined as '"product used to provide electronic information for health and social care purposes".	To pass, the developer is required to confirm that they have undertaken Clinical Risk Management activities in compliance with DCB019.
C1.1.1	Please detail your clinical risk management system	Provided/no evidence available	DCB019 sets out the activities that must and should be undertaken for health IT systems. An example clinical risk management system template can be downloaded[89] from the NHS Digital website.	To pass, the developer is required to evidence that a clinical risk management system is in place and that it is compliant with the requirements set out in DCB019. This should include: – the clinical risk management governance arrangements that are in place; – the clinical risk management activities; – clinical safety competence and training; – audits.

This example of criteria shows the level of detail and the evidence required along with the supporting documentation provided to help developers address each criterion and sub-criterion.

Alongside this list of criteria, the NHS uses a reference framework issued by the *National Institute for Health and Care Excellence* (NICE) known as "*Evidence standards framework for digital health technologies*". An update concerning medical devices was published in April 2021[90].

This reference framework sets out the levels of evidence required based on a functional digital medical device classifications. Note that Nwe (110), in 2020, demonstrated poor reproducibility of this classification (Kappa of 0.32, with 95% CI between 0.16 and 0.47).

As regards the required medical content quality level, the minimum requirements are:

- **valid** (in line with the best available sources, such as NICE guidelines, relevant professional organisations or recognised UK patient associations, relevant to the target population);
- **accurate**;
- **up-to-date**;
- **regularly audited** at defined intervals, for example every year;
- **sufficiently comprehensive**.

In 2020, Rowland (111) listed the functional categories of the NHS approach for patient apps. They include:

- support clinical diagnosis and/or decision making;
- improve clinical outcomes from established treatment pathways through behaviour change and enhancement of patient adherence and compliance with treatment;

[88] https://digital.nhs.uk/data-and-information/information-standards/information-standards-and-data-collections-including-extractions/publications-and-notifications/standards-and-collections/dcb0129-clinical-risk-management-its-application-in-the-manufacture

[89] https://digital.nhs.uk/services/clinical-safety/documentation#clinical-risk-management

[90] https://www.nice.org.uk/corporate/ecd7/chapter/section-a-evidence-for-effectiveness-standards#functional-classification-of-dhts

- act as standalone digital therapeutics;
- primarily to deliver disease related education.

In 2020, around one hundred apps had been indexed in the NHS library.

3.6.4. AppSaludable (Andalusia region, Spain)

Developed by the Andalusia region in 2012, AppSaludable[91] is one of the first platforms to be set up in Europe to catalogue mobile apps (112).

Assessment covers four domains and 31 recommendations, and is performed in four phases (submission form, self-assessment, external assessment, and regional certification).

Details of the recommendations are provided in Appendix 6.

3.6.5. Portugal

Developed on a national scale in Portugal, MySNS Seleção applies the same assessment model as the Andalusia region[92], proposed European guidelines[93], and the European privacy code of conduct[94].

Each app undergoes an assessment process[95] made up of four criteria:

1. **performance**: providing a user-friendly interface, a suitable design enabling access for the greatest number of people, with a view to maximising the advantages offered by technology;
2. **security**: complying with legal data protection requirements, accounting for the sensitive nature of the information recorded and guaranteeing user privacy and confidentiality;
3. **public utility**: ensuring that the app adds value and helps significantly improve individuals' and/or healthcare professionals' lives;
4. **information quality and security**: ensuring information is credible and current, in terms of disclosure, stating information sources, any conflicts of interest, and advertising content.

The assessment process is described in five phases with an extra phase compared to the Andalusia region which is a preparatory phase prior to submitting the app.

[91] http://www.calidadappsalud.com/distintivo/catalogue
[92] https://digitalhealtheurope.eu/twinnings/dhe-twinning-results/appsaludable/
[93] https://ec.europa.eu/health/sites/health/files/ehealth/docs/ev_20160607_co06_04_en.pdf
[94] https://ec.europa.eu/digital-single-market/en/news/code-conduct-privacy-mhealth-apps-has-been-finalised
[95] https://mysns.min-saude.pt/criterios-de-avaliacao/

3.6.6. Summary

Table 8 is a compilation of the examples mentioned above.

Table 8: Examples of organisations providing regional or national editorialised repositories

Name	Country	Assessment system	Number of apps	Comments
NHS Health Apps Library (first launch in 2013-2015 [96] and relaunch in 2017 [beta] [97] and 2019-)[98] updated in 2021	Great Britain	DAQS (from 55 to 243 items) and DAP DAQS: 7 domains reviewed	92 (06/20)	New assessment version in February 2021
NHS Health Apps Library (February 2021)[99]	Great Britain	DTAC 5 domains reviewed		New assessment version (02/21)
AppSaludable[100]	Spain (Andalusia)	4 domains	39 and 81 in progress (01/21)	
MySNS Seleção[101]	Portugal	Twinned with Andalusia	ND	

The number of apps was counted at the end of February 2021. ND: not defined

3.7. Editorialised repository consisting of a database providing a specific annotation of each app in a specialised sector

Some public or private bodies provide app databases (repositories) adding a very in-depth and specific assessment in specific domains.

3.7.1. *Our Mobile Health* (OMH) (Great Britain)

OMH has set up of a database of general[102] or targeted apps, for example relating to Parkinson's disease[103].

The assessment process covers the following domains: data security, regulatory compliance, technical stability, usability, and effectiveness factors. A panel of over 150 independent expert reviewers with clinical and non-clinical backgrounds provides detailed assessments of each app.

OMH has a vast network of health app developers providing content in various fields, particularly diabetes, women's health, fitness, cancer management, nutrition, heart monitoring, and mental health.

The app library contains over 200 apps that have undergone the assessment process.

[96] https://bmcmedicine.biomedcentral.com/articles/10.1186/s12916-015-0451-z
[97] https://mhealthintelligence.com/news/uk-tries-again-with-a-library-of-certified-mobile-health-apps
[98] https://mhealthintelligence.com/news/uks-nhs-moves-to-fast-track-mhealth-app-validation-process
[99] https://www.nhs.uk/apps-library/
[100] http://www.calidadappsalud.com
[101] https://www.mysns.min-saude.pt/mysns-selecao/
[102] https://www.ourmobilehealth.com
[103] https://www.parkinsons.org.uk/information-and-support/apps-and-devices-parkinsons

3.7.2. Mental health with AppChecker (Denmark)

The AppChecker[104] system targets apps in the area of mental health.

The assessment process includes a number of steps:

➔ **Step 1 – App information.**

In this section, the self-assessment documents information covering technical specifications, who the developer is, how the app was developed, the target audience, language, operating system, price, etc.

➔ **Step 2 – Assessment of security and privacy.**

The security and privacy assessment is divided into three steps starting with a risk assessment. The risk assessment includes eight levels (rated from R1 to R8). The app risk level determines the number of parameters to be assessed. If the risk level is greater than R1, a decision tree specifying whether sensitive personal data are collected or not is provided. If the risk level is greater than R5, a decision tree is provided to decide whether the app requires the CE mark.

The three steps for assessing security and privacy:

- Risk assessment. In the risk assessment, a check is run to determine whether the app collects data or not, and if so, which data and how. A risk level is assessed in an assessment table corresponding to the app. The risk level determines whether an app security and privacy check is required and whether it needs to be labelled as a medical device (CE mark).
- Data security and privacy assessment. If the app is ranked above R1 in the risk assessment, security and privacy are assessed. In this section, a decision tree is provided regarding security and privacy, which helps determine whether the app meets security and privacy requirements. Apps and data collection can vary in complexity; as such, some apps will pass the test after one or two items, whereas others will need to go through eight items.
- Assessment of need for CE mark. Few apps will need to comply with the CE mark, but if the risk assessment in step 1 ranks the app in category R5 or over, it is necessary to check whether a CE mark is required.

➔ **Step 3 – Quality assessment.**

This is the final assessment of app quality through 4 domains, each containing three items:

- information and disclosure;
- clinical quality;
- functionality;
- usability.

For each item, a score of 1 to 3 is applied, providing an overall average score. An app must obtain a score greater than 1.50 in order to be recommended.

3.7.3. AppCheck (Germany)

Funded by the Ministry of Health of the Rhineland-North Westphalia region, AppCheck runs an internal assessment on apps for chronic lung conditions (asthma, etc.) and diabetes[105]. It is aimed at helping patients using recommended apps to facilitate disease management.

[104] https://mindapps.dk/en/guidance-to-use-the-app-checker/
[105] https://appcheck.de/zertifizierte-apps-2/

3.7.4. Summary

Table 9 compiles the examples mentioned above.

Table 9: Examples of organisations providing specialised editorialised repositories

Name	Country	Assessment specificities	Number of apps	Comments
AppCheck	Germany	Respiratory medicine and diabetes	17	
AppCheck	Denmark	Mental health	ND	Risk assessment
OMH (*Our Mobile Health*)	Great Britain	Parkinson's	ND	Flexibility on other topics
ORCHA	Great Britain	On request	ND	

The number of apps was counted at the end of February 2021. ND: not defined

3.8. Editorialised repository consisting of a database providing a specific selection of apps defined as medical devices

3.8.1. mHealthBelgium (Belgium)

In Belgium, a portal[106] has been in place since 2019 for mobile apps which are medical devices. Three levels are offered.

Level 1 corresponding to the CE mark, level 2 with an independent risk analysis[107], and level 3 requiring published clinical evidence obtained through randomised controlled trials. This level is reimbursed by Belgian national health insurance.

The descriptive sheet details the criteria for these three tiers[108].

3.8.2. DiGA (Germany)

Germany has set up a medical device assessment process partially focusing on mobile apps, known as DiGA (*Digital Health Applications*)[109]. The **proposed assessment enables preliminary inclusion based on certain criteria.** A one-year probational period is offered to obtain clinical evidence (positive effects of care) leading to referencing in the repository and determination of the clinical benefit.

The process applies the following steps:

- the manufacturer submits their app;
- BfArM (*Bundesinstitut für Arzneimittel und Medizinprodukte*) which is the Federal Institute of Medicines and Medical Devices (equivalent to the French National Agency for Medicines and Health Products Safety – ANSM – in France) issues an opinion and reviews the application;

[106] https://mhealthbelgium.be/fr/toutes-les-apps
[107] https://mhealthbelgium.be/wp-content/uploads/2019/09/Criteria-mHealth-apps-ENv4.pdf
[108] https://mhealthbelgium.be/fr/pyramide-de-validation
[109] https://diga.bfarm.de/de/verzeichnis

- three-month assessment on prerequisites (security, functionality, quality, privacy, data security, interoperability) and on the positive effects on care (medical benefit, structural improvement or procedures);
- there are three possible decisions:
  - rejection;
  - if the prerequisites are met, but not enough evidence is available to demonstrate positive effects, preliminary admission is proposed for a twelve-month period. Justification of plausible findings is required. At the end of this probational period, a rejection or inclusion decision is made;
  - inclusion and publication in the repository (determination of clinical benefit) and potential pricing and reimbursement negotiation.

Similarly to the Belgian repository, the German selection process for **DiGA is also aimed at determining apps eligible for reimbursement based on the clinical benefit**.

3.8.3. Digi-HTA (Finland)

In Finland, a repository also focuses solely on medical devices. Digi-HTA[110] **provides product-specific Digi-HTA recommendations based on company-provided product information**. This information is supplemented by a literature review, an expert assessment, and any additional questions for the company. Product information is compiled using a Digi-HTA questionnaire[111].

In terms of information security and data protection, information is compiled in two documents. These documents are based on the information security and data protection requirements of the Finnish National Cybersecurity Centre (NCSC-FI)[112]. By default, the recommendation is valid for three years. The product will be reassessed if important new information becomes available and the company requests a reassessment.

The assessment covers digital medical devices (mHealth, artificial intelligence, robotics) (113). A total of 11 domains, made up of around one hundred items, are covered, broken down as follows

- company information (3 items);
- product information (21 items);
- technical stability (6 items);
- cost (6 items);
- effectiveness (6 items);
- clinical safety (9 items);
- personal data protection and security (2 separate documents);
- usability and accessibility (11 items and 5 separate documents);
- interoperability (10 items, of which 2 are conditional and 2 separate documents);
- artificial intelligence (22 items);
- robotics (4 items).

A score on a 10-point scale is assigned after assessing five domains scoring 2 points, 1 point or -4 points:

[110] https://www.ppshp.fi/Tutkimus-ja-opetus/FinCCHTA/Sivut/Digi-HTA.aspx
[111] https://www.oulu.fi/cht/digihealthhub/digi-hta
[112] https://www.kyberturvallisuuskeskus.fi/en/ncsc-news/instructions-and-guides/information-security-and-data-protection-requirements-social

- effectiveness (sufficient, promising, poor or unknown);
- safety (sufficient, probably at a sufficient level, but lacking evidence, poor or unknown);
- cost (reasonable, high, unreasonably high);
- data protection and security (sufficient, slightly lacking, lacking);
- usability and accessibility (sufficient, slightly lacking, lacking);

The score is broken down into five levels: 10 (dark green: recommended); 9 (light green: one factor needs to be taken into consideration when using the product); 7-8 (yellow: a number of factors need to be taken into consideration when using the product); 5-6 (orange: a large number of factors need to be taken into consideration using the product); ≥ 4 (red: critical factors need to be taken into consideration when using the product). An exportable document for each product assessed is generated with the score/colour for each domain.

3.8.4. Summary

Table 10 compiles the examples mentioned above.

Table 10: Examples of organisations providing editorialised repositories for medical devices

Name	Country	Assessment system	Number of apps	Comments
mHealthBelgium	Belgium	3 MD tiers with socioeconomic evidence 6 domains	25	Medical device only
DiGA	Germany	2 phases: P1: technical prerequisite + medical effects or improvement of procedures. P2: possible justification.	11	12-month probational period to obtain clinical evidence
Digi-HTA	Finland	11 domains and separate documents	4	Applies to apps and robotics
Evidence standard framework (ESF) for digital health technology [113]	Great Britain	Functional classification and specific items	ND	Based on a functional classification. Update in progress in April 2021

The number of apps was counted at the end of February 2021. ND: not defined

To differentiate the eHealth sector from HTA (*Health Technology Assessment*), digital health products and technologies and assessment processes are categorised under the **term eHTA**, or "*electronic Health Technology Assessment*".

Systematic reviews address issues associated with electronic health technology assessment (HTA).

For example, in relation to:

- assessment approaches (114);
- consistency of *European Network for Health Technology Assessment* (EUnetHTA) (115) criteria on a European level;
- decision-making processes for assessing these digital solutions (116);
- common and specific assessment frameworks for particular products (mHealth, artificial intelligence, robotics) (113) ;

[113] https://www.nice.org.uk/corporate/ecd7

- a specific assessment framework (technical and material, social and personal, political and organisational) (117);
- development of a specific module for apps (118).

The assessment type and duration are a challenge to be tackled for innovative and agile (regularly updated) digital solutions, which are unsuited to overly long processing times or assessment processes. For example, one mental health app disappears every 2.9 days from online stores (119).

The FDA (*Food and Drug Administration*) offers a fast-track route for tools subject to short cycles[114]. The EUnetHTA network also offers "*Rapid Relative Effectiveness Assessments*" (REAs) (114).

However, Alon (120), in 2020, demonstrated that risk-based assessment using FDA categories (pre-certification) did not identify at-risk apps among the top 10 apps covering six health issues (addiction, anxiety, depression, diabetes, high blood pressure, and schizophrenia).

mHealth is a specific section of the digital health sector as a whole and is subject to ever shorter cycles, with frequently limited economic models (free app, etc.).

In conclusion

App **medical content** assessment strategies generally focus on scientific publication quality (source, author expertise, conflicts of interest, currency, levels of evidence). The tools used are scores, self-assessments based on general criteria or specific guidelines for targeted apps.

When the app **uses or processes health data**, assessments become more complex and the assessment tools used range from scores or labels to assessment guidelines of several hundred criteria or external tests.

Organisations (non-profit, commercial or governmental) compile these different assessments and their results in databases in "editorialised repository" format.

In 2020, this overview shows a wide variety of assessment systems within a given country or internationally. Although the assessment domains are the same as a whole, the level of precision of criteria and their breakdown make it difficult to establish equivalence.

European initiatives were implemented in 2016 and are currently being continued with the "*European mHealth Innovation and Knowledge Hub*" or the ISO-82304-2 standard. The objective being to standardise assessment systems so that users and manufacturers can get clearer bearings or set up mutual recognition systems.

Pending any harmonisation, designers must pay particular attention to key information which should be in the public domain and which would enable external app assessment. This information should help comply with the assessment criteria most commonly found in the various assessment tools mentioned above.

[114] https://www.fda.gov/medical-devices/device-software-functions-including-mobile-medical-applications/examples-premarket-submissions-include-mmas-cleared-or-approved-fda

4. Medical content quality criteria

To define medical content quality criteria, this document is based on the 2016 HAS guidelines (1) published to assess mHealth apps. At the time, it set out five domains and 14 subdomains which are listed below:

- user information (description, consent);
- **health content** (initial content design, standardisation, generated content, interpreted content);
- technical content holder (technical design, data stream);
- security/trustworthiness (cybersecurity, trustworthiness, confidentiality);
- usability/use (usability/design, acceptability, integration/import).

4.1. Different types of health content

In terms of health content, the guidelines made a distinction between content published to disseminate general information (known as initial content) and user-generated content *via* data collected by smart objects, questionnaires, built-in biosensors in the smartphone, etc.

For this document, which only relates to medical content, this distinction is also used, with **four different sections** corresponding to specific types of content:

- **initial content**: linked with the inform or instruct function in general;
- **generated content**: linked with user-generated data;
- **interpreted content**: linked with data produced by the user and interpreted by a professional or an algorithm;
- **displayed content**: linked with the manner in which data are published by the creator and understood by the user.

4.2. Use of medical content quality criteria

In the following subsections, a selection of criteria has been made based on the 2016 HAS guidelines (1) and after an external review (selection process described in Appendix 7). The final selection of the 17 criteria selected is given below.

Each criterion is justified and accompanied with examples of documents or information to be provided.

The number of criteria tailored to the app assessment is linked with the end-use of the app. For example, for an app collecting no data, only the criteria associated with initial content and displayed content need to be entered. The other criteria associated with data collection are irrelevant.

This selection can be linked up with an app's **basic functions** described above: inform, instruct, record, calculate/analyse, alert/remind, display, guide, and communicate (see section 2).
Furthermore, the risk level associated with app use could be linked with the three aspects described above:

- the **end-use** of the app claimed by the manufacturer (risks associated with the claimed objectives of the app, risks associated with delayed care, etc.);
- the **user profile** (risks for a vulnerable population, novice user, etc.);
- the **context of use** (risks in the data collection processing, unsuitable use setting, etc.).

The assessment should be documented in proportion to the severity of the identified risk.

4.3. Initial content design

These criteria are to be applied for any health-related data production.

4.3.1. Information management

The expected requirement level depends on the risk of providing incomplete, incorrect, outdated, or biased information for different reasons or causes.

Information management quality is expressed by the trustworthiness of published information.

- ➔ Is the published information clearly established reference information or is it subject to controversy or disagreement?
- ➔ What are the potential risks and their severity level in the event of poor information management on the topic addressed by the app?

4.3.1.1. Information service organisation

Criterion: presence of an organisation or an editorial expert committee selecting and validating the information published in the app.

Justification: the writing and management of the content available in the app are based on an organisation which validates and guarantees the quality of the information published.

Type of information to be provided: specify the members of the committee which must be made up of experts covering the field with either the relevant CV or the expert's credentials. List links to published information. In the case of good practice guidelines, list links to source documents.

4.3.1.2. Expertise of authors of content in app

Criterion: experts (healthcare professionals, engineers, algorithm experts, professional bodies, patient or consumer associations, etc.) are involved in producing the content available in the app.

Justification: the level of expertise of the authors of the content of the app is a sign of quality. Recognition by peers or endorsement by professional bodies or associations improves the credibility of the published content.

Type of information to be provided: list where the name of the authors of the content and their references or credentials can be viewed online, or are freely accessible. In the case of good practice guidelines, only list links to source documents.

4.3.1.3. Declarations of interest

Criterion: declarations of conflicts of interest of the different contributors are available to view for all.

Justification: declarations of any conflicts of interest are a sign of transparency for users and external assessors. Conflicts of interest can result in bias which might call into question the trustworthiness of the product.

Type of information to be provided: list where conflicts of interest can be consulted online, or are freely accessible. In the case of good practice guidelines, only list links to source documents.

4.3.1.4. Quote of key sources and bibliographic references

Criterion: key sources and references for publications justifying app content are documented and are available to view by all.

Justification: in the health sector, quoting bibliographic sources and an objective selection of the best data available is a requisite sign of quality. Access to the reference list must be easy to view.

Type of information to be provided: describe how these sources and references are presented. The reference may be quoted within the app, or on a resource website, or by external documentation, or as a reference at the end of the content, etc. In the case of good practice guidelines, links to source documents are listed.

4.3.1.5. Update of key sources and bibliographic references

Criterion: the process for monitoring and updating key sources and references in relation to publications is documented.

Justification: bibliographic monitoring helps update and adapt the current knowledge processed by the app. The date of information update is to be cited.

Type of information to be provided: describe where the date of information update is published and how updating is carried out to the user.

4.3.1.6. Level of evidence

Criterion: where a specific assessment of the product and levels of evidence is available, these suitable references are available to view by all.

Justification: the HAS has issued guidance documents for critical literature review, grading levels of evidence, or assessment methodology[115,116 ,117 ,118 ,119].

Some apps have been the subject of a randomised controlled trial (RCT) or some types of apps have been the subject of systematic reviews. These references are key and must be accessible to justify the benefit of the product.

Type of information to be provided: state how these levels of evidence are presented.

[115] http://www.has-sante.fr/portail/upload/docs/application/pdf/analiterat.pdf
[116] http://www.has-sante.fr/portail/upload/docs/application/pdf/2013-06/etat_des_lieux_niveau_preuve_gradation.pdf
[117] http://www.has-sante.fr/portail/upload/docs/application/forcedownload/2016-03/guide_methodologique_analyse_critique.pdf
[118] http://www.has-sante.fr/portail/upload/docs/application/pdf/eval_interventions_ameliorer_pratiques_guide.pdf
[119] http://www.has-sante.fr/portail/upload/docs/application/pdf/2011-11/guide_methodo_vf.pdf

4.4. Generated content and standardisation

If the app does not generate data, the following criteria are not applicable.

4.4.1. Relevance, trustworthiness of information and real-life justification

The expected requirement level depends on the risk associated with poor information collection in real-life scenarios.

App use tracking quality is expressed by feedback management and product maintenance.

- ➔ Is the type of data enough to fulfil the object(s) set by the app?
- ➔ Does the type of data refer to precision, reproducibility, granularity and accuracy characteristics.
- ➔ What is the app follow-up policy under real-life conditions?
- ➔ What are the potential risks and their level of severity in event of a lack of follow-up?

4.4.1.1. Measurement quality in context of use

Criterion: the measurement quality (contextual robustness) in the setting or context of use is documented and justified with regard to the end-use of the product.

Justification: the measurement made under real-life conditions may differ from measurements made in a laboratory setting.

The measurement of health or wellness data in the user's setting must be in line with the objective of the app.

Type of information to be provided: provide documentation specifying measurement performance under real-life conditions. This follow-up can be illustrated by concrete examples.

4.4.1.2. User support

Criterion: support is provided and allows users to request assistance for product-use related queries (understanding content and using features). Frequently asked questions are documented and updated.

Justification: product use support helps improve quality of use. This support can adopt different forms depending on the objectives of the products and different formats of use.

Type of information to be provided: list and describe the user support systems provided and tailored to the user (hotline, FAQ, moderated discussion threads, user forums with charter of use, etc.).

4.5. Interpreted content

If the app does not generate data, the following criteria are not applicable.

The expected requirement level depends on the risk of providing information interpreted incorrectly for different causes.

Interpretation quality is expressed by the trustworthiness of the information analysed and analysis capabilities.

- ➔ Is the interpretation policy clearly defined?
- ➔ What are the potential risks and their severity level in the event of poor management of the interpretation of the information targeted by the app?

4.5.1. Interpretation typology

Some apps provide content interpretation. The interpretation can be carried out by a professional who is an expert in the field or assisted by an algorithm. Regarding the use of artificial intelligence (AI), Haverinen (113) proposed in 2019 a 20-item list to be applied for AI use. In 2021, a European regulation is under development in relation to AI[120]. A specifically adapted assessment will shortly be implemented based on this new regulation.

According to current knowledge, it is necessary to ensure that these tools can refer to a healthcare professional in the event of doubt.

4.5.1.1. Human interpretation of health content

Criterion: in the case of human (non-automated) interpretation of content for health purposes (health data, scientific content, etc.), this is carried out by suitable healthcare professionals for the topic studied or specifically trained competent persons.

Justification: the interpretation of scientific content or health data requires the involvement of qualified and competent persons.

Type of information to be provided: describe who, when (first-line, second-line, etc.) and how interpretation is implemented.

4.5.1.2. Automated interpretation of health content

This criterion is liable to change over the coming years, as the mHealth sector and the current technology available help develop algorithms and AI (see above).

Criterion: in the event of interpretation by algorithms intended, for example, to interpret content with a medical remit or for advisory purposes, the user must be informed whether or not AI is included in the algorithm(s) used. The exhaustiveness of the studies assessing the performance level of algorithms should be available to the user.

Justification: automated interpretation of scientific content or health data requires assessment of the trustworthiness of the interpretation. The credibility of algorithm tests is a critical factor to be assessed to ensure trustworthiness.

Type of information to be provided: list all publications supporting the use of the algorithms in the app. If specific publications are unavailable, explain how the algorithms are used and how the tests to ensure interpretation performance are carried out.

[120] https://eur-lex.europa.eu/procedure/FI/2021_106

4.6. Displayed content

The expected requirement level depends on the risk of providing displayed information which might be difficult for the user to understand.

4.6.1. User understanding and involvement

The expected requirement level depends on the risk of failing to meet the expectations of future users of the app.

Requirements, user tests, or assessment of the degree of understanding of app use by different user profiles are quality factors which enhance the product.

- ➔ What was the role of users in app design and validation?
- ➔ What are the potential risks and their level of severity associated with a lack of user involvement?

4.6.1.1. Involvement of users (patients, professionals, specific parties)

Criterion: the main users are involved in the different app development phases.

Justification: design with the different stakeholders, specified transparently, is a sign of quality.

Type of information to be provided: describe the strategy for involving users in app development. Concrete examples may illustrate the impact and benefit of this involvement (review grid, writing tool, *Living Lab*, etc.).

4.6.1.2. Description of end-use

Criterion: the main end-use (objective or purpose) of the product is the subject of a precise description available to view by all.

Justification: this statement is an important factor for defining the intended use of the app.

If the use declared by the manufacturer is an instrument, apparatus, appliance or software to be used for human beings for, in particular, diagnosis, prevention, monitoring, treatment, alleviation of an illness or an injury (Regulation [EU] 2017/745 of the European Parliament and of the Council of 5 April 2017 on medical devices), it is eligible to be a medical device. The manufacturer could contact the French National Agency for Medicines and Health Products Safety (ANSM) and comply with the applicable legal and regulatory provisions[121].

Type of information to be provided: specify where the primary objective of the product is described. Specify how the creator unambiguously explains what the app does and does not do.

4.6.1.3. Contraindications, risks, limitations of use

Criterion: contraindications, risks of incorrect interpretation and limitations of use are assessed and documented. This information is available for users to access and view.

Justification: the app may have limitations of use or of trustworthiness.

Type of information to be provided: specify limitations of use or of trustworthiness that the app may have. Specify where and how the creator publishes this information transparently, comprehensibly and tailored to the user.

[121] http://ansm.sante.fr/Produits-de-sante/Dispositifs-medicaux

4.6.1.4. Understanding of health content

Criterion: the app uses terminology that is understandable for the user.

Justification: ambiguous terms, mistranslation or the use of jargon may result in poor understanding or interpretation of the content provided.

Type of information to be provided: specify the creator's strategy for making the information provided accessible and understandable (external review, user test, thesaurus, tooltip, etc.).

4.6.2. Content display performance

As the user is generally autonomous when reading the display content, it is necessary to ensure that the information generated is clearly understood.

- Is the display clear enough, comprehensible and suitable?
- The display of content involves sensitive issues in relation to ergonomics (e.g.: the information should be entirely visible without being cut off on the screen, etc.), the use of icons, for example, in a data security level alert, when reading blocks of text on the screen, etc.
- It also involves accounting for the context of use and users, with a view to anticipating their perception of the information and their ability to understand it with or without user help tools.
- What are the potential risks and their severity levels in the event of poor management of the display of information by the app?

4.6.2.1. User help/instructions

Criterion: a user help system for the product is provided to users (contextual help, online help, user manual, tutorial, teaching software, e-learning, etc.). This system supports the user's ability to learn.

Justification: the level of teaching support provided by the creator to the user helps optimise use of the product.

Type of information to be provided: list the resources used to aid app use.

4.6.2.2. Text and image readability and navigation

Criterion: the navigation and readability of the various media used (text, image, videos) have been tested. The interface allows changes to app readability (change of font or font size, alert colour, etc.).

Justification: the navigation and readability of information by users of different abilities are a factor in product accessibility.

Type of information to be provided: describe the strategy used to help optimise navigation and text and image readability for the user.

4.6.2.3. Error prevention and understanding of information

Criterion: a tailored alert system for critical decisions following any errors in understanding information by the user is set up to prevent risks or redirect to a healthcare professional.

Justification: some interpretation errors can give rise to decision-making errors by the user. Alert information should be guaranteed by the creator along with explanations or descriptions of the reasons for the alert generated.

Type of information to be provided: describe the strategy for preventing information interpretation errors by the user.

Table of appendices

Annexe 1. Referencing in the Digital health space (ENS) and the Professional service package (BSP) 60

Annexe 2. The Information Standard for Health and Care Information Production Quality Statements (NHS – UK) 62

Annexe 3. *Mobile App Rating Scale* (MARS) 64

Annexe 4. ORCHA-24 (*Organisation for the Review of Care and Health Applications*) 66

Annexe 5. Example of detailed "clinical content" for assessing pain management apps (Reynoldson, 2014 (95)) 67

Annexe 6. Details of AppSaludable regional assessment recommendations (Spain, Andalusia region) 68

Annexe 7. Criterion selection process 70

Annexe 8. Literature search 75

Table of figures

Figure 1. Evolution of the number of publications in the mHealth sector from 2005 to 2020 16

Table of tables

Table 1: Basic functions of an app and examples of risks 22

Table 2: Examples of scores solely assessing medical content 29

Table 3: Examples of composite scores assessing different mHealth domains 33

Table 4: Examples of multipurpose guidelines for assessing mHealth 34

Table 5: Examples of websites featuring user and/or expert app reviews 37

Table 6: Examples of organisations providing labels, certification or accreditation 40

Table 7: Examples of criteria described in NHS guidelines (Great Britain) 44

Table 8: Examples of organisations providing regional or national editorialised repositories 46

Table 9: Examples of organisations providing specialised editorialised repositories 48

Table 10: Examples of organisations providing editorialised repositories for medical devices 50

Table 11: Comparison of the domains and subdomains included between 2016 and 2021 (pre-review) 70

Table 12: Comparison of the criteria included between 2016 and 2021 (pre-review) 71

Table 13: Comparison of criteria included after review by review group 73

Annexe 1. Referencing in the Digital health space (ENS) and the Professional service package (BSP)

Mon espace santé corresponds to action 16 of the digital health roadmap[122]. The digital service package for healthcare professionals (BSP) corresponds to action 17 of the roadmap.

The Ministerial delegation for digital healthcare requested that the HAS contribute to the drafting of the referencing framework by setting out the minimum base in respect of medical content quality criteria. A decree setting out the referencing procedure and the referencing committee shall be published in the first half of 2021.

This work is being conducted in conjunction with the Delegation for digital health (DNS), in concert with the National health insurance fund for salaried workers (CNAM) leading the ENS/BSP taskforce and the French digital healthcare agency (ANS) with a trial version in the second half of 2021 and a public version available from January 2022.

The following digital services are eligible:

- mobile apps;
- software;
- web platforms.

The eligible digital services are aimed at users (patients/individuals) and are published by public or private actors, in the healthcare (health and wellness), medico-social and social sectors. They must fall within the scope defined by Article L 1111-13-1 of the French Public Health Code. They may be:

- optionally linked to one or more smart objects;
- free or fee-based;
- medical devices or not.

The candidate service for referencing must comply with all the legal and regulatory, national and community provisions in force. This particularly applies to services falling within the remit of European regulation 2017/745 on medical devices which must meet the relevant safety and performance requirements. Compliance with these regulations is expressed by the CE medical device mark.

Each candidate service for referencing is assessed with respect to:

- technical digital health governance (ethical, safety, interoperability guidelines);
- requirements in relation to medical content quality criteria;
- criteria in relation to personal data protection;
- requirements in relation to interaction with the ENS space.

Compliance with technical digital health governance and compliance with requirements in relation to content quality criteria and certain requirements in relation to personal data protection are documented *via* the Convergence platform[123].

A Referencing Committee is co-chaired by the DNS and CNAM to deliver an opinion for referencing.

A summary of the assessment of each candidate service for referencing is presented to the Committee.

[122] https://esante.gouv.fr/virage-numerique/feuille-de-route
[123] https://convergence.esante.gouv.fr

The Committee's opinions can take two forms:

- an unfavourable opinion for referencing, which must be justified (explaining the reasons). The publisher is then invited to correct the service according to a non-conformity processing procedure;
- a favourable opinion for referencing.

The favourable opinion may be accompanied by feedback with a view to improving the service.

The decision is then taken by the Minister of Health. In the event of a decision in favour of referencing, a three-party agreement is drawn up between the publisher, the Minister of Health, and CNAM, for each service referenced.

The decision is notified to the publisher in its private space in the publisher portal.

Annexe 2. The Information Standard for Health and Care Information Production Quality Statements (NHS – UK)

The NHS has set out a list of six principles and 16 items for assessing the information production process. These principles are devised to demonstrate that the manufacturer has implemented a process for producing and maintaining high-quality, evidence-based, health and care-related information. This information should be available for the user.

Information based on the website is given below[124] (the questionnaire is aimed at the organisation producing the information).

Information production

This principle is designed to demonstrate that your organisation has a defined process in place to produce good quality health and care information in a consistent manner.

- Quality statements:
  - There is a defined process for producing information (including identifying the need for a product, checking stages, final sign-off, review, version control and archiving);
  - all individuals involved in the information process have the relevant up to date training/experience and follow the defined process for all information products.

Evidence sources

This principle is designed to ensure that where evidence is used, is it relevant and from a recognised source.

- Quality statements:
  - information is created using high quality evidence (where the evidence exists) and is presented in a balanced manner. Where there is no evidence to back up claims made in an information product this is made clear to the end user. Evidence is reviewed for currency each time the resource is updated;
  - information is reviewed by relevant professionals/peers before it is approved for use.

User understanding and involvement

This principle is designed to ensure through user involvement that you understand who the information is designed for, why it is required and what users' needs are. And to ensure, through user testing, that it reflects those needs and the views of those using it. User involvement, including testing, should be representative of the target audience and involve an appropriate number of such people.

- Quality statements:
  - information is created taking into consideration the health literacy and/or accessibility needs of the population it is aimed at;
  - jargon is not used and medical terms (when used) are explained;
  - end users are involved at the outset and throughout in the production and their input is actively used.

[124] https://www.england.nhs.uk/tis/about/the-info-standard/#information-production

End product

This principle is designed to ensure that the product has been developed following your process and is of high quality.

- Quality statements:
  - an authorised approver(s) checks that your process has been followed sufficiently before a product is approved for publication;
  - the date and review date of each information product are clearly stated;
  - information is in plain language, free from spelling and grammatical errors – and medical terms are explained where necessary;
  - references to the evidence used in the information are retained and made available if requested;
  - the information signposts the end user to further sources of information;
  - the information product gives the end user details on how they can give their feedback.

Feedback

This principle is designed to ensure that all feedback (outside the development process) is dealt with appropriately especially concerning errors, omissions or points for clarification. Such feedback is recorded, actioned and resolved as appropriate, especially if an amendment to, or withdrawal of, an information product is required.

- Quality statements:
  - people are encouraged to give any ongoing feedback after the product has been published and this is acted upon as appropriate.

Review

This principle is designed to ensure that your information products are reviewed on a planned and regular basis, within a timeframe appropriate to the type of information, not normally more than every three years. Any products that are not reviewed within your defined review periods should no longer be distributed.

- Quality statements:
  - there is a defined process for reviewing published/approved information;
  - all staff involved in the information process follow the defined review process for all information products.

Annexe 3. *Mobile App Rating Scale* (MARS)

The MARS score was produced by the *Queensland University of Technology* (121) under the author Stoyanov (75).

The following is based on an unvalidated French translation.

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

Score out of 25

- Entertainment: Is the app fun/entertaining to use? Does it use any strategies to increase engagement through entertainment (e.g. through gamification)? (Items are rated on a 5-point scale with descriptors)
- Interest: Is the app interesting to use? Does it use any strategies to increase engagement by presenting its content in an interesting way? (Items are rated on a 5-point scale with descriptors)
- Customisation: Does it provide/retain all necessary settings/preferences for apps features (e.g. sound, content, notifications, etc.)? (Items are rated on a 5-point scale with descriptors)
- 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. (Items are rated on a 5-point scale with descriptors)
- Target group: Is the app content (visual information, language, design) appropriate for your

  target audience? (Items are rated on a 5-point scale with descriptors)

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

Score out of 20

- Performance: How accurately/fast do the app features (functions) and components

  (buttons/menus) work? (Items are rated on a 5-point scale with descriptors)
- Ease of use: How easy is it to learn how to use the app; how clear are the menu labels/icons and instructions? (Items are rated on a 5-point scale with descriptors)
- Navigation: Is moving between screens logical/accurate/appropriate/ uninterrupted; are all necessary screen links present? (Items are rated on a 5-point scale with descriptors)
- Gestural design: Are interactions (taps/swipes/pinches/scrolls) consistent and intuitive across all components/screens? (Items are rated on a 5-point scale with descriptors)

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

Score out of 15

- Layout: Is arrangement and size of buttons/icons/menus/content on the screen appropriate or zoomable if needed? (Items are rated on a 5-point scale with descriptors)
- Graphics: How high is the quality/resolution of graphics used for buttons/icons/menus/content? (Items are rated on a 5-point scale with descriptors)
- Visual appeal: How good does the app look? (Items are rated on a 5-point scale with descriptors)

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.

Score out of 35

- Accuracy of app description (in app store): Does app contain what is described? (Items are rated on a 5-point scale with descriptors)
- Goals: Does app have specific, measurable and achievable goals? (Items are rated on a 5-point scale with descriptors)
- Quality of information: Is app content correct, well written, and relevant to the goal/topic of the app? (Items are rated on a 5-point scale with descriptors)
- Quantity of information: Is the extent coverage within the scope of the app; and comprehensive but concise? (Items are rated on a 5-point scale with descriptors)
- Visual information: Is visual explanation of concepts – through charts/graphs/images/videos, etc. – clear, logical, correct? (Items are rated on a 5-point scale with descriptors)
- Credibility: Does the app come from a legitimate source (specified in app store description or within the app itself)? (Items are rated on a 5-point scale with descriptors)
- Evidence base: Has the app been trialled/tested; must be verified by evidence (in published scientific literature)? (Items are rated on a 5-point scale with descriptors)

Total quality score: A+B+C+D

Subjective section

Section E

Score out of 20

- Would you recommend this app to people who might benefit from it? (Items are rated on a 5-point scale with descriptors)
- How many times do you think you would use this app in the next 12 months if it was relevant to you? (Items are rated on a 5-point scale with descriptors)
- Would you pay for this app? (Items are rated on a 5-point scale with descriptors)
- What is your overall star rating of the app? (Items are rated on a 5-point scale with descriptors)

Annexe 4. ORCHA-24 (*Organisation for the Review of Care and Health Applications*)

Short version of the ORCHA questionnaire published by Leigh in 2018 (87)

Three categories are listed, representing 24 items and a maximum score of 24. The section on quality assurance is detailed below.

- Data governance (8 questions)
- Clinical efficacy and assurance (7 questions):
  - Is there a statement within the app itself, or the app store, about user feedback during design, development or testing?
  - Is there a statement either in the app or store about user involvement in testing?
  - Is there a statement within the app that it has been tested and shown to be beneficial to someone with the relevant condition?
  - Is there a statement within the app, or app store, about the app having been through a clinical trial, or other form of testing to show real world effectiveness, and has received positive feedback?
  - Is there a statement about how frequently any advice, guidance or content will be reviewed to ensure accuracy and clinical relevance?
  - Is there a statement within the app that it has been positively evaluated or validated by a clinical or other relevant expert?
  - Is there any evidence within the app that the developer has attempted to validate any guidance or recommendations with academic expertise?
  - Is there a statement within the app identifying a list of review or accrediting bodies or individuals?
- User experience and engagement (8 questions).

Annexe 5. Example of detailed "clinical content" for assessing pain management apps (Reynoldson, 2014 (95))

A good application for recording pain should include the ability to input the following details where appropriate: (Y/N)

- Site: Where is the pain?
- Onset:
  - What date did the pain occur?
  - What time did the pain occur?
- Character: What is the pain like? What patterns does it appear in?
- Radiation: Does it move anywhere else?
- Associated symptoms: Is there anything else with it (e.g., nausea, blurred vision, and numbness)?
- Timing: How long was the duration of the pain episode?
- Exacerbating or relieving factors:
  - Does anything make it worse?
  - Does anything make it better?
- Environment: Place in which the pain occurred (e.g., home, work, and outdoors);
- Severity: A rating scale (e.g., numerical—from 1 to 10, with 10 being worst pain);
- Medication or treatment: A separate treatment section allowing users to input details of medication taken or treatment undergone:
  - Does it include such a section?
  - What medication was taken/treatment undergone?
  - When did it happen?
  - How much was taken? (dose of medication; details of treatment)
  - How did it affect the pain? Did it make it better or worse, or have no effect?
- Notes: Ability to add extra notes about the episode (e.g., if something was important but not covered by the questions).
- Can options be modified (e.g., options can be added or removed to make application more personal/relevant)? (Y/N)
- Does the application remain true to the claims made by the developer in the market description?
- Results:
  - Is there a results section? (Y/N)
  - Can results be filtered? (Y/N)
  - Does it generate reports? (Y/N).

Annexe 6. Details of AppSaludable regional assessment recommendations (Spain, Andalusia region)

Developed by the Andalusia region in 2012, AppSaludable[125] is one of the first platforms to be set up in Europe to catalogue mobile apps (112).

Assessment covers four domains and 31 recommendations, and is performed in four phases (submission form, self-assessment, external assessment, and regional certification). Details are provided in Appendix 5.

Details of the domains and recommendations are listed below.

Design and appropriateness

- **Appropriateness**. Recommendation 1. The health App clearly defines its functional reach and its purpose, identifying the target groups of information and the aims pursued regarding these groups.
- **Accessibility**. Recommendation 2. The health App follows the Principles of Universal Design, as well as reference accessibility standards and recommendations.
- **Design**. Recommendation 3. The health App follows the recommendations, patterns and directives on matters of design included in the official manuals of the different platforms.
- **Usability/Testing**. Recommendation 4. The health App has been tested by potential users before its availability to the public.

Quality and safety of information

- **Suitability for the audience**. Recommendation 5. The health App adapts itself to its target audience.
- **Transparency**. Recommendation 6. The health App offers transparent information about its owners' identity and location. Recommendation 7. The health App offers information about its funding sources, promotion and sponsorship, as well as about possible conflicts of interests.
- **Authorship**. Recommendation 8. The health App identifies the authors of its content, as well as their professional qualification.
- **Information update/revisions**. Recommendation 9. The health App includes the date of the last revision made in the published material. Recommendation 10. The health App warns of those updates which modify or influence the functioning of health-related content, as well as other sensitive data.
- **Content and information sources**. Recommendation 11. The health App is based on one or more reliable information sources, and takes into account the available scientific evidence. Recommendation 12. The health App offers concise information about the procedure used in order to select its content. Recommendation 13. The health App is based on ethical principles and values.
- **Risk management**. Recommendation 14. The possible risks for patient safety caused by the use of the health App are identified. Recommendation 15. The known risks and adverse events (near misses) are analysed, and the convenient actions start to be developed.

Provision of services

[125] http://www.calidadappsalud.com/distintivo/catalogue

- **Technical support/Inquiries**. Recommendation 16. The health App has a support system about its use. Recommendation 17. The health App offers a contact mechanism for technical support with an assured and fixed response time.
- **E-Commerce**. Recommendation 18. The health App informs about the terms and conditions on its products and services' commercialisation.
- **Bandwidth**. Recommendation 19. The health App makes an efficient use of communications bandwidth.
- **Advertisement**. Recommendation 20. The health App warns of the use of advertisement mechanisms and allows deactivating or skipping it.

Confidentiality and privacy

- **Privacy and data protection**. Recommendation 21. Before downloading and installing, the health App informs about the kind of user's data to be collected and the reason, about the access policies and data treatment, and about possible commercial agreements with third parties. Recommendation 22. The health App clearly describes the terms and conditions about recorded personal data. Recommendation 23. The functioning of the health App preserves privacy in the recorded information collects express consents granted by users and warns of risks coming from the use of online mobile health Apps. Recommendation 24. The health App ensures pertinent security measures when users' health information or sensitive data has to be collected or exchanged. Recommendation 25. The health App informs the users when it has access to other resources of the device, to users' accounts and to profiles in social networks. Recommendation 26. The health App ensures the right of access to recorded information and the updates regarding changes in its privacy policy. Recommendation 27. The health App has measures regarding minors' protection in accordance with the current legislation.
- **Logical security**. Recommendation 28. The health App neither presents any sort of known susceptibility nor any type of malicious code. Recommendation 29. The health App describes the security procedures established in order to avoid unauthorised access to personal data collected, as well as to limit the access by third parties. Recommendation 30. The health App has encryption mechanisms for the storage and exchange of information, as well as mechanisms for passwords management. Recommendation 31. When the health App uses services from the Cloud (cloud computing), the terms and conditions of those services are declared, and the pertinent security measures are ensured.

Annexe 7. Criterion selection process

The 2016 HAS guidelines (1) contained 101 criteria divided into five different domains.

The purpose of this work was to only sect **quality criteria for assessing health content**.

The "user information", "technical content holder" and "security/trustworthiness" are not within the scope of the 2021 request, with the sole exception of the "contraindications, risks, limitations of use" criterion which featured in the "trustworthiness" subdomain and has been included in the "initial content" domain.

The "health content" domain and some of the criteria of the "usability/use" domain were included before submitting for review (Table 11).

Table 11: Comparison of the domains and subdomains included between 2016 and 2021 (pre-review)

2016 domains and subdomains	2021 pre-review domains
User information: – description – consent	Not included in relation to the requested topic. They are covered by specific technical criteria for the ENS.
Health content: – initial content design – standardisation – generated content – interpreted content	– initial content design – generated content and standardisation – interpreted content The criteria of the "standardisation" domain are partially included. The criteria in relation to the "interoperability" format as a whole are covered by specific technical criteria for the ENS and are not included in this work.
Technical content holder: – technical design – data stream	Not included in relation to the requested topic. They are covered by specific technical criteria for the ENS.
Security/Trustworthiness: – cybersecurity – trustworthiness – privacy	Not included in relation to the requested topic, with the exception of the "**Contraindications, risks, limitations of use**" criterion from the trustworthiness subdomain which is retained and incorporated with the other initial content criteria. They are covered by specific technical criteria for the ENS.
Usability/use: – **usability/design** – acceptability – integration/import	– displayed content (from "usability/design") The two remaining subdomains are not included in relation to the requested topic. They are covered by specific technical criteria for the ENS.

In a more detailed way, selecting the "health content" domain published in 2016 made it possible to include all of the criteria published across the three types of content (initial, general, interpreted). The "standardisation" subdomain was partially included, as most of the technical criteria (particularly in respect of interoperability) are coved specifically elsewhere[126].

The content may also be interpreted incorrectly when displayed, which has led to the creation of a "display content" domain which is a summary of the 2016 "usability/design" subdomain (Table 12). It encompasses the criteria relating to the format for interpreting displayed content.

In total, 28 criteria of the 101 criteria were retain prior to submission to the review group.

[126] https://convergence.esante.gouv.fr

Table 12: Comparison of the criteria included between 2016 and 2021 (pre-review)

2016 subdomains and criteria	2021 domains/subdomains and criteria (pre-review)
Initial content design: – involvement of users (patients, professionals, specific parties) – user requirement engineering methodology – information service organisation – content authors' expertise – declarations of interest – quotes of key sources and bibliographic references – update of key sources and bibliographic references – level of evidence – description of end-use – product language – thesaurus-glossary	**Initial health content/information management**: – information service organisation – content authors' expertise – declarations of interest – quotes of key sources and bibliographic references – update of key sources and bibliographic references – level of evidence – Initial health content/user understanding and involvement: – involvement of users (patients, professionals, specific parties) – user requirement engineering methodology – description of end-use – contraindications, risks, limitations of use – product language – thesaurus-glossary
Standardisation: – interoperability: semantic standards, reference terminologies – data accuracy and reproducibility – data granularity – information loss (by aggregation, by compression, etc.) – measurement performance in context of use – data synchronisation capability	**Generated content and standardisation/trustworthiness of data**: – data accuracy and reproducibility – information loss (by aggregation, by compression, etc.) The other criteria do not fall within the remit of the topic covered. They are covered by specific technical criteria for the ENS.
Generated content: – relevance of data collected – minimisation of data collected – number of interfaces/devices/apps – relevance of information in context – electronic discussion threads – operational support, hotline	**Generated content and standardisation/information relevance and support under real-life conditions**: – measurement performance in context of use – relevance of information in context – electronic discussion threads – operational support, hotline The other criteria do not fall within the remit of the topic covered. They are covered by specific technical criteria for the ENS.
Interpreted content: – algorithm types – human interpretation of health content – automated interpretation of health content	– **Interpreted content/interpretation typology**: – algorithm types – human interpretation of health content – automated interpretation of health content
Usability/design: – ergonomics – installation process and setup – user help/instructions – usability and user-friendliness – text and image readability – level of use – content accessibility for people with disabilities – ease of use – error prevention – use cases, business scenarios – flexibility/customisation – response times, display times	**Displayed content/content display performance**: – ergonomics – user help/instructions – usability and user-friendliness – text and image readability – Displayed content/content adaptation to user: – level of use – error prevention – use cases, business scenarios The other criteria do not fall within the remit of the topic covered. They are covered by specific technical criteria for the ENS.

The review group received the 28 criteria and the document reviewing the literature. An assessment as regards substance, format, and supporting data was requested (Table 13).

Regarding initial content, the substance-related feedback concerned trade secrecy-related issues for disseminating information from the user tests conducted (engineering of needs) and problems obtaining objective assessment data in this area.

Regarding the language used for the product or tools such as the thesaurus, the feedback noted that the requirement level and utility of these criteria were dependent on the type of app. They were merged and replaced by “understanding of health content”.

Regarding generated content, the criteria relating to measurement accuracy and reproducibility were deemed useful, but, for some complex apps where there was no gold standard for example or based on patient profiles. These criteria were grouped together under a single “measurement performance in context of use” criterion.

Regarding the hotline criterion or discussion threads, the feedback pointed out that the cost of these services and the fact that these services may not be suitable for some apps. A “user support” criterion makes it possible to provide tailored solutions without any specific ones being mentioned as a criterion.

Regarding interpreted content, the criterion relating to information on the “type” of algorithm used was deemed somewhat irrelevant, due to being too complex for the user to understand.

Regarding displayed content, the feedback pointed out that the proposed criteria were either removed from the “health content” theme, or already redundant with the “user involvement” criterion assessing the interface, or complex to objectify in a pragmatic way. They were grouped together into three criteria: user help/understanding all content and an alert system for displays causing potential errors.

Table 13: Comparison of criteria included after review by review group

Pre-review (28 criteria)	After review by review group (17 criteria)
Initial health content/information management: – information service organisation – content authors' expertise – declarations of interest – quotes of key sources and bibliographic references – update of key sources and bibliographic references – level of evidence Initial health content/user understanding and involvement: – involvement of users (patients, professionals, specific parties) – user requirement engineering methodology – description of end-use – contraindications, potential risks, limitations of use – product language – thesaurus-glossary	– Initial health content/information management: – information service organisation – expertise of authors of content in app – declarations of interest – quotes of key sources and bibliographic references – update of key sources and bibliographic references – level of evidence The "initial health content/user understanding and involvement" section has been moved to the displayed content level. As user involvement applies to all types of content
Generated content and standardisation/trustworthiness of data: – data accuracy and reproducibility – information loss (by aggregation, by compression, etc.)	
Generated content and standardisation/information relevance and support under real-life conditions: – measurement performance in context of use – relevance of information in context – electronic discussion threads – operational support, hotline	Generated content and standardisation/information relevance, trustworthiness and support under real-life conditions: – measurement quality in context of use – user support
Interpreted content/interpretation typology: – algorithm types – human interpretation of health content – automated interpretation of health content	Interpreted content/interpretation typology: – human interpretation of health content – automated interpretation of health content
Displayed content/content display performance: – ergonomics – user help/instructions – usability and user-friendliness – text and image readability Displayed content/content adaptation to user: – level of use – error prevention – use cases, business scenarios	Initial health content/user understanding and involvement: – involvement of users (patients, professionals, specific parties) – description of end-use – contraindications, potential risks, limitations of use – understanding of health content Displayed content/content display performance: – user help/instructions – text and image readability and navigation – error prevention

In total, 17 criteria were retained or are obtained from merging the 28 criteria submitted to the review group (Table 13).

In terms of format, minor edits were made to improve the description of criterion expectations.

Regarding the evidence to be provided to be able to assess the criterion, suggestions were made to make the assessment more realistic and clearer by the creator. The creator is given some leeway based on the type of apps to be assessed.

According to the apps and their potential risk level, the degree of accuracy of this evidence requires an assessment by healthcare professional or competent persons to review its relevance.

Annexe 8. Literature search

The documentary search, restricted to publications in English and French, was conducted on the Medline database. The database query strategy specifies for each question and/or study type the search terms used, the Boolean operators, and the search period.

The search terms are either thesaurus terms (descriptors), or free-text terms (from the title or the abstract). They are combined with the terms describing the study types.

Table 1 shows a summarised version of the steps of this query in the Medline database. The total number of references obtained by querying this bibliographic database is 2073.

Table 1: Medline database search strategy

Study type/subject		Period
	Terms used	
Guidelines and consensus conferences		01/2015 – 12/2020
Step 1	Mobile Applications/de OR (mobile application OR mobile applications OR mobile app OR mobile apps OR smartphone application OR smartphone applications OR Smartphone app OR Smartphone apps OR app stores OR Mobile Medical Application OR Mobile Medical Applications OR medical apps OR medical app OR standalone software OR health apps OR health app OR mhealth OR mobile health OR ehealth OR apps OR app)/ti OR ((Medical Informatics Applications/de OR Software/majr OR application/ti OR medical/ti) AND (mobile OR smartphone OR phone)/ti)	
AND		
Step 2	(recommendation* OR guideline* OR statement* OR consensus OR position paper)/ti OR health planning guidelines/de OR (practice guideline OR guideline OR Consensus Development Conference OR Consensus Development Conference, NIH)/pt	
Meta-analyses & systematic reviews		01/2015 – 12/2020
Step 1		
AND		
Step 3	(metaanalys* OR meta-analys* OR meta-analysis OR systematic review* OR systematic overview* OR systematic literature review* OR systematical review* OR systematical overview* OR systematical literature review* OR systematic literature search OR pooled analysis)/ti OR (meta-analysis OR "Systematic Review")/pt OR cochrane database syst rev/sd	
App assessment		01/2015 – 12/2020

Step 1		
AND		
Step 4	(framework OR frameworks OR certificat* OR label* OR standard OR criteria OR scoring OR quality assessment)/ti OR (Mobile App Rating scale OR MARS)/ti,ab OR Software Validation/de	
OR		
Step 5	(mhealth OR mobile health OR apps OR app OR standalone software)/ti AND (evaluat* OR Assessment)/ti	
Scales and apps		01/2015 – 01/2021
Step 1		
AND		
Step 6	(scale OR reliabilit* OR validit* OR assessment)/ti OR (scale development) OR (app ability)/ab)	

*: truncation; de: descriptor; ti: title; ab: abstract; pt: publication type; so: journal title

Results

Number of references identified: 2073

Numbers of references reviewed: 483

Number of references selected: 121

Bibliographic references

1. Haute Autorité de Santé. Référentiel de bonnes pratiques sur les applications et les objets connectés en santé (mobile Health ou mHealth). Saint-Denis La Plaine: HAS; 2016. https://www.has-sante.fr/upload/docs/application/pdf/2016-11/has_ref_apps_oc.pdf

2. Pagliari C, Sloan D, Gregor P, Sullivan F, Detmer D, Kahan JP, et al. What is eHealth (4): a scoping exercise to map the field. J Med Internet Res 2005;7(1):e9. http://dx.doi.org/10.2196/jmir.7.1.e9

3. Webb TL, Joseph J, Yardley L, Michie S. Using the internet to promote health behavior change: a systematic review and meta-analysis of the impact of theoretical basis, use of behavior change techniques, and mode of delivery on efficacy. J Med Internet Res 2010;12(1):e4. http://dx.doi.org/10.2196/jmir.1376

4. Salmani H, Ahmadi M, Shahrokhi N. The impact of mobile health on cancer screening: A systematic review. Cancer Inform 2020;19:1-11. http://dx.doi.org/10.1177/1176935120954191

5. Leppin AL, Montori VM, Gionfriddo MR. Minimally disruptive medicine: a pragmatically comprehensive model for delivering care to patients with multiple chronic conditions. Healthcare 2015;3(1):50-63. http://dx.doi.org/10.3390/healthcare3010050

6. Wallace D, Perry J, Yu J, Mehta J, Hunter P, Cross KM. Assessing the Need for Mobile Health (mHealth) in Monitoring the Diabetic Lower Extremity. JMIR Mhealth Uhealth 2019;7(4):e11879. http://dx.doi.org/10.2196/11879

7. Scott K, Richards D, Adhikar R. A review and comparative analysis of security risks and safety measures of mobile health apps. Australas J Inf Syst 2015;19:1-17. http://dx.doi.org/10.3127/ajis.v19i0.1210

8. Kumar S, Nilsen WJ, Abernethy A, Atienza A, Patrick K, Pavel M, et al. Mobile health technology evaluation: the mHealth evidence workshop. Am J Prev Med 2013;45(2):228-36. http://dx.doi.org/10.1016/j.amepre.2013.03.017

9. IQVIA Institute for Human Data Science. The Growing value of digital health. Parsippany: IQVIA Institute; 2017. https://regresearchnetwork.org/wp-content/uploads/the-growing-value-of-digital-health.pdf

10. Fagherazzi G, Ravaud P. Digital diabetes: Perspectives for diabetes prevention, management and research. Diabetes Metab 2019;45(4):322-9. http://dx.doi.org/10.1016/j.diabet.2018.08.012

11. PricewaterhouseCoopers. Socio-economic impact of mHealth. An assessment report for the European Union. Ahmedabad: PwC; 2013. http://www.gsma.com/connectedliving/wp-content/uploads/2013/06/Socio-economic_impact-of-mHealth_EU_14062013V2.pdf

12. Powell A, Torous J. A patient-centered framework for measuring the economic value of the clinical benefits of digital health apps: Theoretical modeling. JMIR Ment Health 2020;7(10):e18812. http://dx.doi.org/10.2196/18812

13. Wicks P, Chiauzzi E. 'Trust but verify'--five approaches to ensure safe medical apps. BMC Med 2015;13:205. http://dx.doi.org/10.1186/s12916-015-0451-z

14. Dick S, O'Connor Y, Thompson MJ, O'Donoghue J, Hardy V, Wu TJ, et al. Considerations for Improved Mobile Health Evaluation: Retrospective Qualitative Investigation. JMIR Mhealth Uhealth 2020;8(1):e12424. http://dx.doi.org/10.2196/12424

15. Albrecht UV. Créer la transparence et fournir une orientation. Méthodes et outils d'aide décisionnelle pour l'utilisation des applications mobiles de santé. Bern: eHealth Suisse; 2019. https://mhh-publikationsserver.gbv.de/receive/mhh_mods_00000003

16. Albrecht UV, Hillebrand U, von Jan U. Relevance of trust marks and CE labels in German-language store descriptions of health apps: Analysis. JMIR Mhealth Uhealth 2018;6(4):e10394. http://dx.doi.org/10.2196/10394

17. World Health Organization. Monitoring and evaluating digital health interventions: a practical guide to conducting research and assessment. Geneva: WHO; 2016. https://apps.who.int/iris/bitstream/handle/10665/252183/9789241511766-eng.pdf

18. World Health Organization. WHO guideline: recommendations on digital interventions for health system strengthening. Geneva: WHO; 2019.

19. Pham Q, Wiljer D, Cafazzo JA. Beyond the randomized controlled trial: a review of alternatives in mhealth clinical trial methods. JMIR Mhealth Uhealth 2016;4(3):e107. http://dx.doi.org/10.2196/mhealth.5720

20. Byambasuren O, Sanders S, Beller E, Glasziou P. Prescribable mHealth apps identified from an overview of systematic reviews. NPJ Digit Med 2018;1:12. http://dx.doi.org/10.1038/s41746-018-0021-9

21. Mohr DC, Cheung K, Schueller SM, Hendricks Brown C, Duan N. Continuous evaluation of evolving behavioral intervention technologies. Am J Prev Med 2013;45(4):517-23. http://dx.doi.org/10.1016/j.amepre.2013.06.006

22. Grundy QH, Wang Z, Bero LA. Challenges in assessing mobile health app quality: A systematic review of prevalent and innovative methods. Am J Prev Med 2016;51(6):1051-9. http://dx.doi.org/10.1016/j.amepre.2016.07.009

23. Ologeanu-Taddei R. Assessment of mhealth interventions: Need for new studies, methods, and guidelines for study designs. JMIR Med Inform 2020;8(11):e21874. http://dx.doi.org/10.2196/21874

24. Huckvale K, Morrison C, Ouyang J, Ghaghda A, Car J. The evolution of mobile apps for asthma: an updated systematic assessment of content and tools. BMC Med 2015;13:58. http://dx.doi.org/10.1186/s12916-015-0303-x

25. Asthana S, Jones R, Sheaff R. Why does the NHS struggle to adopt eHealth innovations? A review of macro, meso and micro factors. BMC Health Serv Res 2019;19(1):984.

http://dx.doi.org/10.1186/s12913-019-4790-x

26. Wierda E, Blok S, Somsen GA, van der Velde ET, Tulevski II, Stavrov B, et al. Protecting patient privacy in digital health technology: the Dutch m-Health infrastructure of Hartwacht as a learning case. BMJ Innov 2020;6:170-6.

27. Zelmer J, van Hoof K, Notarianni M, van Mierlo T, Schellenberg M, Tannenbaum C. An assessment framework for e-mental health apps in canada: Results of a modified delphi process. JMIR Mhealth Uhealth 2018;6(7):e10016. http://dx.doi.org/10.2196/10016

28. Peng W, Kanthawala S, Yuan S, Hussain SA. A qualitative study of user perceptions of mobile health apps. BMC Public Health 2016;16(1):1158. http://dx.doi.org/10.1186/s12889-016-3808-0

29. Pustozerov E, von Jan U, Albrecht UV. Evaluation of mHealth applications quality based on user ratings. Stud Health Technol Inform 2016;226:237-40.

30. Singh K, Drouin K, Newmark LP, Lee J, Faxvaag A, Rozenblum R, et al. Many mobile health apps target high-need, high-cost populations, but gaps remain. Health Aff 2016;35(12):2310-8. http://dx.doi.org/10.1377/hlthaff.2016.0578

31. Ali EE, Chew L, Yap KY-L. Evolution and current status of mhealth research: a systematic review. BMJ Innov 2016;2:33-40.

32. Fiordelli M, Diviani N, Schulz PJ. Mapping mHealth research: a decade of evolution. J Med Internet Res 2013;15(5):e95. http://dx.doi.org/10.2196/jmir.2430

33. Marcolino MS, Oliveira JAQ, D'Agostino M, Ribeiro AL, Alkmim MBM, Novillo-Ortiz D. The Impact of mHealth Interventions: Systematic Review of Systematic Reviews. JMIR Mhealth Uhealth 2018;6(1):e23. http://dx.doi.org/10.2196/mhealth.8873

34. Grainger R, Devan H, Sangelaji B, Hay-Smith J. Issues in reporting of systematic review methods in health app-focused reviews: A scoping review. Health Informatics J 2020:1460458220952917. http://dx.doi.org/10.1177/1460458220952917

35. World Health Organization. Rapid risk assessment of acute public health events. Geneva: WHO; 2012. https://apps.who.int/iris/bitstream/handle/10665/70810/WHO_HSE_GAR_ARO_2012.1_eng.pdf?sequence=1

36. Lewis TL, Wyatt JC. mHealth and mobile medical Apps: a framework to assess risk and promote safer use. J Med Internet Res 2014;16(9):e210. http://dx.doi.org/10.2196/jmir.3133

37. Bradway M, Carrion C, Vallespin B, Saadatfard O, Puigdomènech E, Espallargues M, et al. mHealth Assessment: Conceptualization of a Global Framework. JMIR Mhealth Uhealth 2017;5(5):e60. http://dx.doi.org/10.2196/mhealth.7291

38. Rasmussen J. Risk management in a dynamic society: a modelling problem. Saf Sci 1997;27:183-213.

39. Neumann K, Larisch K, Dietzel J, Kurepkat M, Wenzlau V. Digital healthcare products. Leveraging opportunities – developing safe routes to market. Berlin: IGES Institute; 2016. https://www.iges.com/e6/e1621/e10211/e13470/e15278/e15279/e15281/attr_objs15282/IGES_Publication_Digital_healthcare_products_WEB_eng.pdf

40. Therapeutic Goods Administration. Actual and potential harm caused by medical software. A rapid literature review of safety and performance issues. Woden: TGA; 2020. https://www.tga.gov.au/sites/default/files/actual-and-potential-harm-caused-medical-software.pdf

41. Huckvale K, Prieto JT, Tilney M, Benghozi PJ, Car J. Unaddressed privacy risks in accredited health and wellness apps: a cross-sectional systematic assessment. BMC Med 2015;13:214. http://dx.doi.org/10.1186/s12916-015-0444-y

42. Kearney AT. Mobile Health: Fata Morgana oder Wachstumstreiber? Düsseldorf: AT Kearney GmbH; 2013.

43. Benjumea J, Ropero J, Rivera-Romero O, Dorronzoro-Zubiete E, Carrasco A. Privacy Assessment in Mobile Health Apps: Scoping Review. JMIR Mhealth Uhealth 2020;8(7):e18868. http://dx.doi.org/10.2196/18868

44. Timmers T, Janssen L, Kool RB, Kremer JA. Educating patients by providing timely information using smartphone and tablet apps: Systematic review. J Med Internet Res 2020;22(4):e17342. http://dx.doi.org/10.2196/17342

45. Schmidt-Kraepelin M, Toussaint PA, Thiebes S, Hamari J, Sunyaev A. Archetypes of gamification: Analysis of mHealth Apps. JMIR Mhealth Uhealth 2020;8(10):e19280.

46. Thimbleby H, Oladimeji P, Cairns P. Unreliable numbers: error and harm induced by bad design can be reduced by better design. Journal of the Royal Society, Interface 2015;12(110):0685. http://dx.doi.org/10.1098/rsif.2015.0685

47. Council of Europe. Guidelines on artificial intelligence and data protection. Strasbourg: CE; 2019. https://rm.coe.int/guidelines-on-artificial-intelligence-and-data-protection/168091f9d8

48. Huckvale K, Car M, Morrison C, Car J. Apps for asthma self-management: a systematic assessment of content and tools. BMC Med 2012;10:144. http://dx.doi.org/10.1186/1741-7015-10-144

49. Huckvale K, Adomaviciute S, Prieto JT, Leow MK, Car J. Smartphone apps for calculating insulin dose: a systematic assessment. BMC Med 2015;13:106. http://dx.doi.org/10.1186/s12916-015-0314-7

50. Bierbrier R, Lo V, Wu RC. Evaluation of the accuracy of smartphone medical calculation apps. J Med Internet Res 2014;16(2):e32. http://dx.doi.org/10.2196/jmir.3062

51. Suso-Ribera C, Castilla D, Zaragozá I, Mesas Á, Server A, Medel J, et al. Telemonitoring in chronic pain management using smartphone apps: A randomized controlled trial comparing usual assessment against app-based monitoring with and without clinical alarms. Int J Environ Res Public Health 2020;17(18). http://dx.doi.org/10.3390/ijerph17186568

52. Lewis TL, Boissaud-Cooke MA, Aungst TD, Eysenbach G. Consensus on use of the term "App" versus "Application" for reporting of mHealth research. J Med Internet Res 2014;16(7):e174; discussion e. http://dx.doi.org/10.2196/jmir.3460

53. Riezebos RJ. Peer-reviewing of mHealth applications. Requirements for peer-reviewing mobile health applications and development of an online peer review tool Amsterdam: University of Amsterdam; 2014. http://dare.uva.nl/cgi/arno/show.cgi?fid=573074

54. Eysenbach G. CONSORT-EHEALTH: improving and standardizing evaluation reports of Web-based and mobile health interventions. J Med Internet Res 2011;13(4):e126. http://dx.doi.org/10.2196/jmir.1923

55. Agarwal S, LeFevre AE, Lee J, L'Engle K, Mehl G, Sinha C, et al. Guidelines for reporting of health interventions using mobile phones: mobile health (mHealth) evidence reporting and assessment (mERA) checklist. BMJ 2016;352:i1174. http://dx.doi.org/10.1136/bmj.i1174

56. Boudreaux ED, Waring ME, Hayes RB, Sadasivam RS, Mullen S, Pagoto S. Evaluating and selecting mobile health apps: strategies for healthcare providers and healthcare organizations. Transl Behav Med 2014;4(4):363-71. http://dx.doi.org/10.1007/s13142-014-0293-9

57. Guidance on evaluating or developing a health app. Wellington: Ministry of Health Ministry of Health; 2017. https://www.health.govt.nz/system/files/documents/pages/guidance-evaluating-developing-health-app-oct17-v2.pdf

58. Nielsen SL, Rimpiläinen S. T1.2.1 Report on international practice on digital apps. T1 Technology : Digital Health & Care Institute; 2018.

59. Institut für Gesundheitsförderung und Prävention GmbH, Kraus-Füreder H. Gesundheits Apps. Grundlagenpapier unter besonderer Berücksichtigung des Aspekts Gesundheitskompetenz. Graz: IFGP ; 2018.

60. Nouri R, S RNK, Ghazisaeedi M, Marchand G, Yasini M. Criteria for assessing the quality of mHealth apps: a systematic review. Journal of the American Medical Informatics Association : JAMIA 2018;25(8):1089-98. http://dx.doi.org/10.1093/jamia/ocy050

61. Urmet H. Mapping the nice evidence standards framework and international experiences into a guideline for the assessment of mobile health solutions in Estonia [Thesis] Tallinn: University of technology School of Information Technologies.

62. European mHealth hub. D2.1 knowledge tool 1. health apps assessment frameworks (Pending EC approval). Seville: mHealth-hub; 2020. https://mhealth-hub.org/download/d2-1-knowledge-tool-1-health-apps-assessment-frameworks-pending-ec-approval

63. Llorens-Vernet P, Miró J. The Mobile App Development and Assessment Guide (MAG): Delphi-Based Validity Study. JMIR Mhealth Uhealth 2020;8(7):e17760. http://dx.doi.org/10.2196/17760

64. Levine DM, Co Z, Newmark LP, Groisser AR, Holmgren AJ, Haas JS, et al. Design and testing of a mobile health application rating tool. NPJ Digit Med 2020;3:74. http://dx.doi.org/10.1038/s41746-020-0268-9

65. Albrecht UV. Chances and risks of mobile health apps (CHARISMHA). Hannover: Hannover Medical School; 2016. https://publikationsserver.tu-braunschweig.de/receive/dbbs_mods_00060024

66. Azad-Khaneghah P, Neubauer N, Miguel Cruz A, Liu L. Mobile health app usability and quality rating scales: a systematic review. Disabil Rehabil Assist Technol 2020:1-10. http://dx.doi.org/10.1080/17483107.2019.1701103

67. Haute Autorité de Santé. Évaluation de la qualité des sites e-santé et de la qualité de l'information de santé diffusée sur Internet (Revue de la littérature des outils d'évaluation). Saint-Denis La Plaine: HAS; 2007. https://www.has-sante.fr/upload/docs/application/pdf/evaluation_qualite_site_sante_internet.pdf

68. Van Singer M, Chatton A, Khazaal Y. Quality of smartphone apps related to panic disorder. Front Psychiatry 2015;6:96. http://dx.doi.org/10.3389/fpsyt.2015.00096

69. Butcher R, MacKinnon M, Gadd K, LeBlanc-Duchin D. Development and examination of a rubric for evaluating point-of-care medical applications for mobile devices. Med Ref Serv Q 2015;34(1):75-87. http://dx.doi.org/10.1080/02763869.2015.986794

70. Silberg WM, Lundberg GD, Musacchio RA. Assessing, controlling, and assuring the quality of medical information on the Internet: Caveant lector et viewor--Let the reader and viewer beware. JAMA 1997;277(15):1244-5.

71. Zhang MW, Ho RC, Loh A, Wing T, Wynne O, Chan SWC, et al. Current status of postnatal depression smartphone applications available on application stores: an information quality analysis. BMJ Open 2017;7(11):e015655. http://dx.doi.org/10.1136/bmjopen-2016-015655

72. Jeon E, Park HA, Min YH, Kim HY. Analysis of the information quality of korean obesity-management smartphone applications. Healthc Inform Res 2014;20(1):23-9. http://dx.doi.org/10.4258/hir.2014.20.1.23

73. Khazaal Y, Chatton A, Cochand S, Coquard O, Fernandez S, Khan R, et al. Brief DISCERN, six questions for the evaluation of evidence-based content of health-related websites. Patient Educ Couns 2009;77(1):33-7. http://dx.doi.org/10.1016/j.pec.2009.02.016

74. Guan Z, Sun L, Xiao Q, Wang Y. Constructing an assessment framework for the quality of asthma smartphone applications. BMC Med Inform Decis Mak 2019;19(1):192. http://dx.doi.org/10.1186/s12911-019-0923-8

75. Stoyanov SR, Hides L, Kavanagh DJ, Zelenko O, Tjondronegoro D, Mani M. Mobile app rating scale: a new tool for assessing the quality of health mobile apps. JMIR Mhealth Uhealth 2015;3(1):e27. http://dx.doi.org/10.2196/mhealth.3422

76. Domnich A, Arata L, Amicizia D, Signori A, Patrick B, Stoyanov S, et al. Development and validation of the Italian version of the Mobile Application Rating Scale and its generalisability to apps targeting primary prevention. BMC Med Inform Decis Mak 2016;16:83. http://dx.doi.org/10.1186/s12911-016-0323-2

77. Martin Payo R, Fernandez Álvarez MM, Blanco Díaz M, Cuesta Izquierdo M, Stoyanov SR, Llaneza Suárez E. Spanish adaptation and validation of the Mobile Application Rating Scale questionnaire. Int J Med Inform 2019;129:95-9. http://dx.doi.org/10.1016/j.ijmedinf.2019.06.005

78. Messner EM, Terhorst Y, Barke A, Baumeister H, Stoyanov S, Hides L, et al. The German version of the mobile app rating scale (MARS-G): Development and validation study. JMIR Mhealth Uhealth 2020;8(3):e14479.

http://dx.doi.org/10.2196/14479

79. Bardus M, Awada N, Ghandour LA, Fares EJ, Gherbal T, Al-Zanati T, et al. The arabic version of the mobile app rating scale: development and validation study. JMIR Mhealth Uhealth 2020;8(3):e16956.
http://dx.doi.org/10.2196/16956

80. Terhorst Y, Philippi P, Sander LB, Schultchen D, Paganini S, Bardus M, et al. Validation of the Mobile Application Rating Scale (MARS). PloS one 2020;15(11):e0241480.
http://dx.doi.org/10.1371/journal.pone.0241480

81. Terhorst Y, Messner EM, Schultchen D, Paganini S, Portenhauser A, Eder AS, et al. Systematic evaluation of content and quality of English and German pain apps in European app stores. Internet Interv 2021;24:100376.
http://dx.doi.org/10.1016/j.invent.2021.100376

82. Salazar A, de Sola H, Failde I, Moral-Munoz JA. Measuring the quality of mobile apps for the management of pain: Systematic search and evaluation using the mobile app rating scale. JMIR Mhealth Uhealth 2018;6(10):e10718.
http://dx.doi.org/10.2196/10718

83. Dawson RM, Felder TM, Donevant SB, McDonnell KK, Card EB, 3rd, King CC, et al. What makes a good health 'app'? Identifying the strengths and limitations of existing mobile application evaluation tools. Nurs Inq 2020;27(2):e12333.
http://dx.doi.org/10.1111/nin.12333

84. Stach M, Kraft R, Probst T, Messner EM, Terhorst Y, Baumeister H, et al. Mobile health app database - A repository for quality ratings of mhealth apps. Dans: 2020 IEEE 33rd International Symposium on Computer-Based Medical Systems (CBMS) 2020. 2020. 427-32.

85. Stoyanov SR, Hides L, Kavanagh DJ, Wilson H. Development and validation of the User Version of the Mobile Application Rating Scale (uMARS). JMIR Mhealth Uhealth 2016;4(2):e72.
http://dx.doi.org/10.2196/mhealth.5849

86. Wyatt JC, Thimbleby H, Rastall P, Hoogewerf J, Wooldridge D, Williams J. What makes a good clinical app? Introducing the RCP Health Informatics Unit checklist. Clin Med 2015;15(6):519-21.
http://dx.doi.org/10.7861/clinmedicine.15-6-519

87. Leigh S, Ouyang J, Mimnagh C. Effective? Engaging? Secure? Applying the ORCHA-24 framework to evaluate apps for chronic insomnia disorder. Evid Based Ment Health 2017;20(4):e20.
http://dx.doi.org/10.1136/eb-2017-102751

88. Jin M, Kim J. Development and evaluation of an evaluation tool for healthcare smartphone applications. Telemed J E-Health 2015;21(10):831-7.
http://dx.doi.org/10.1089/tmj.2014.0151

89. Jake-Schoffman DE, Silfee VJ, Waring ME, Boudreaux ED, Sadasivam RS, Mullen SP, et al. Methods for evaluating the content, usability, and efficacy of commercial mobile health apps. JMIR Mhealth Uhealth 2017;5(12):e190.
http://dx.doi.org/10.2196/mhealth.8758

90. Modave F, Bian J, Leavitt T, Bromwell J, Harris Iii C, Vincent H. Low quality of free coaching apps with respect to the American College of Sports Medicine guidelines: A review of current mobile apps. JMIR Mhealth Uhealth 2015;3(3):e77.
http://dx.doi.org/10.2196/mhealth.4669

91. Guo Y, Bian J, Leavitt T, Vincent HK, Vander Zalm L, Teurlings TL, et al. Assessing the Quality of Mobile Exercise Apps Based on the American College of Sports Medicine Guidelines: A Reliable and Valid Scoring Instrument. J Med Internet Res 2017;19(3):e67.
http://dx.doi.org/10.2196/jmir.6976

92. Choi J, Noh GY, Park DJ. Smoking cessation apps for smartphones: content analysis with the self-determination theory. J Med Internet Res 2014;16(2):e44.
http://dx.doi.org/10.2196/jmir.3061

93. Ouhbi S, Fernández-Alemán JL, Pozo JR, Bajta ME, Toval A, Idri A. Compliance of blood donation apps with mobile OS usability guidelines. J Med Syst 2015;39(6):63.
http://dx.doi.org/10.1007/s10916-015-0243-1

94. van Galen LS, Xu X, Koh MJA, Thng S, Car J. Eczema apps conformance with clinical guidelines: a systematic assessment of functions, tools and content. Br J Dermatol 2020;182(2):444-53.
http://dx.doi.org/10.1111/bjd.18152

95. Reynoldson C, Stones C, Allsop M, Gardner P, Bennett MI, Closs SJ, et al. Assessing the quality and usability of smartphone apps for pain self-management. Pain Med 2014;15(6):898-909.
http://dx.doi.org/10.1111/pme.12327

96. Myint M, Adam A, Herath S, Smith G. Mobile phone applications in management of enuresis: The good, the bad, and the unreliable! J Pediatr Urol 2016;12(2):112 e1-6.
http://dx.doi.org/10.1016/j.jpurol.2015.09.011

97. Middelweerd A, Mollee JS, van der Wal CN, Brug J, Te Velde SJ. Apps to promote physical activity among adults: a review and content analysis. Int J Behav Nutr Phys Act 2014;11:97.
http://dx.doi.org/10.1186/s12966-014-0097-9

98. McMillan B, Hickey E, Patel MG, Mitchell C. Quality assessment of a sample of mobile app-based health behavior change interventions using a tool based on the National Institute of Health and Care Excellence behavior change guidance. Patient Educ Couns 2015.
http://dx.doi.org/10.1016/j.pec.2015.10.023

99. Siebert JN, Lacroix L, Cantais A, Manzano S, Ehrler F. The impact of a tablet app on adherence to american heart association guidelines during simulated pediatric cardiopulmonary resuscitation: Randomized controlled trial. J Med Internet Res 2020;22(5):e17792.
http://dx.doi.org/10.2196/17792

100. Watson P, Watson D, Dhesi A, New HV, Davidson A, Armstrong R, et al. Improving blood-prescribing decisions: Evaluating the efficacy of a blood guidelines app. Transfus Med 2020.
http://dx.doi.org/10.1111/tme.12736

101. DiFilippo KN, Huang W, Chapman-Novakofski KM. A new tool for nutrition App Quality Evaluation (AQEL): Development, validation, and reliability testing. JMIR Mhealth Uhealth 2017;5(10):e163.
http://dx.doi.org/10.2196/mhealth.7441

102. Anderson K, Burford O, Emmerton L. App chronic disease checklist: Protocol to evaluate mobile apps for chronic disease self-management. JMIR Res Protoc 2016;5(4):e204.
http://dx.doi.org/10.2196/resprot.6194

103. Hanraha C, Dy Aungst T, Cole S. Evaluating mobile medical applications [En ligne]: ASHP eReports; 2014. https://www.ashp.org/-/media/store%20files/mobile-medical-apps.pdf

104. Loy JS, Ali EE, Yap KY. Quality assessment of medical apps that target medication-related problems. J Manag Care Spec Pharm 2016;22(10):1124-40. http://dx.doi.org/10.18553/jmcp.2016.22.10.1124

105. Camacho E, Hoffman L, Lagan S, Rodriguez-Villa E, Rauseo-Ricupero N, Wisniewski H, et al. Technology Evaluation and Assessment Criteria for Health Apps (TEACH-Apps): Pilot Study. J Med Internet Res 2020;22(8):e18346. http://dx.doi.org/10.2196/18346

106. Henson P, David G, Albright K, Torous J. Deriving a practical framework for the evaluation of health apps. Lancet Digit Health 2019;1(2):e52-e4. http://dx.doi.org/10.1016/s2589-7500(19)30013-5

107. van Haasteren A, Vayena E, Powell J. The mobile health app trustworthiness checklist: Usability assessment. JMIR Mhealth Uhealth 2020;8(7):e16844. http://dx.doi.org/10.2196/16844

108. Kaltheuner M, Drossel D, Heinemann L. DiaDigital Apps: Evaluation of smartphone apps using a quality rating methodology for use by patients and diabetologists in Germany. J Diabetes Sci Technol 2019;13(4):756-62. http://dx.doi.org/10.1177/1932296818803098

109. Albrecht UV. Catalogue de critères uniformes pour l'autodéclaration de la qualité des applications de santé. Bern: eHealth Suisse; 2019. https://mhh-publikationsserver.gbv.de/receive/mhh_mods_00000279

110. Nwe K, Larsen ME, Nelissen N, Wong DC. Medical mobile app classification using the national institute for health and care excellence evidence standards framework for digital health technologies: Interrater reliability study. J Med Internet Res 2020;22(6):e17457. http://dx.doi.org/10.2196/17457

111. Rowland SP, Fitzgerald JE, Holme T, Powell J, McGregor A. What is the clinical value of mHealth for patients? NPJ Digit Med 2020;3:4. http://dx.doi.org/10.1038/s41746-019-0206-x

112. Gessa A, Jiménez A, Sancha P. Open innovation in digital healthcare: Users' discrimination between certified and non-certified mhealth applications. J Open Innov Technol Mark Complex 2020;6(4):130.

113. Haverinen J, Keränen N, Falkenbach P, Maijala A, Kolehmainen T, Reponen J. Digi-HTA: Health technology assessment framework for digital healthcare services. FinJeHeW 2019;11:326-41.

114. Vis C, Bührmann L, Riper H, Ossebaard HC. Health technology assessment frameworks for eHealth: A systematic review. Int J Technol Assess Health Care 2020;36(3):204-16. http://dx.doi.org/10.1017/s026646232000015x

115. Kolasa K, Kozinski G. How to Value Digital Health Interventions? A Systematic Literature Review. Int J Environ Res Public Health 2020;17(6). http://dx.doi.org/10.3390/ijerph17062119

116. Oortwijn W, Determann D, Schiffers K, Tan SS, van der Tuin J. Towards integrated health technology assessment for improving decision making in selected countries. Value Health 2017;20(8):1121-30. http://dx.doi.org/10.1016/j.jval.2017.03.011

117. Jacob C, Sanchez-Vazquez A, Ivory C. Understanding Clinicians' Adoption of Mobile Health Tools: A Qualitative Review of the Most Used Frameworks. JMIR Mhealth Uhealth 2020;8(7):e18072. http://dx.doi.org/10.2196/18072

118. Moshi MR, Tooher R, Merlin T. Development of a health technology assessment module for evaluating mobile medical applications. Int J Technol Assess Health Care 2020;36(3):252-61. http://dx.doi.org/10.1017/s0266462320000288

119. Larsen ME, Nicholas J, Christensen H. Quantifying app store dynamics: Longitudinal tracking of mental health apps. JMIR Mhealth Uhealth 2016;4(3):e96. http://dx.doi.org/10.2196/mhealth.6020

120. Alon N, Stern AD, Torous J. Assessing the food and drug administration's risk-based framework for software precertification with top health apps in the united states: Quality improvement study. JMIR Mhealth Uhealth 2020;8(10):e20482. http://dx.doi.org/10.2196/20482

121. Queensland University of Technology. Mobile Application Rating Scale (MARS). App classification. Brisbane: QUT; 2015.

Participants

The participants in the drafting of the 2016 HAS guidelines (1) were requested to review this work with contributions from new participants.

Review group

- Professor Marie-Christine Bene, haematologist, Nantes
- Mr Jean-Luc Bernard, user representative, Pontault-Combault
- Mr Olivier Blot, engineer, CEO of Santexium Consulting, Clermont-Ferrand
- Ms Élisabeth Bouchez, senior medico-technical manager, Créteil
- Ms Anne Buisson, user representative, Paris
- Dr François Carbonnel, general practitioner, Cabestany, lecturer, Montpellier
- Dr Pascal Charbonnel, general practitioner, Paris
- Dr Didier Cugy, general practitioner, Bordeaux
- Professor Fabrice Denis, medical oncologist, Le Mans
- Mr Guillaume de Durat, healthcare law course supervisor, member of the digital healthcare council ethics team, Paris
- Professor Patrick Disdier, internal medicine, Marseille
- Mr Alexandre Falzon, engineer, CEO of Nouveal e-santé, Irigny
- Dr Matthieu Faure, PhD in computer science, engineer, Nimes
- Professor Sébastien Faure, pharmacist, Angers
- Doctor Sylvia Franc, endocrinologist, Corbeil-Essones
- Doctor Jean-Louis Fraysse, pharmacist, member of the digital healthcare council ethics team, Toulouse
- Ms Claire-Charlotte Gaulier, paediatric nurse, Strasbourg
- Dr Leïla Gofti-Laroche, hospital pharmacist, PhD in epidemiology, Grenoble
- Professor Sinem-Ezgi Gülmez, medical doctor and pharmacist, Istanbul
- Mr Marin Guy, physiotherapist, Pessac
- Mr Philippe Haïk, research engineer, Paris
- Dr Olivier Heloir, pharmacist, Lille
- Professor Bruno Housset, respiratory medicine specialist, Créteil
- Dr Cécile Hubsch, neurologist, Paris
- Dr Benjamin Kretz, vascular surgeon, Colmar
- Ms Lyra Ladurée, nurse, healthcare manager, Paris
- Dr Loïck Menvielle, professor, researcher in management science, EDHEC, Nice
- Mr Jean-David Miquel, CEO of Monitorlinq France, Paris
- Dr Alexandre Morel, public health, Brest
- Mr Jean-Philippe Nanteuil, director of innovative digital solutions, Paris
- Professor Gregory Ninot, academic, Montpellier
- Mr Serge Payeur, computer engineer, Rochefort
- Dr Gregory Perrard, cardiologist, Bailleul
- Ms Tatiana Rak, legal manager, Hyères
- Mr Philippe Ruyer, physiotherapist, Les Angles
- Mr Bernard Sauvestre, user representative, Évry-Courcourones
- Dr Éric Sermet, psychiatrist, Châtillon-en-Michaille
- Dr Catherine Sztal-Kutas, user representative, La Rochelle
- Ms Soumia Taoui, clinical research project manager, Toulouse
- Mr Romain Tavignot, nurse anaesthetist, Cannes

Acknowledgements

The HAS would like to thank all the participants listed above for their contribution.

5G	Fifth generation
ACDC	App Chronic Disease Checklist
AI	Artificial intelligence
ANS	Agence du numérique en santé (French digital healthcare agency)
ANSM	Agence Nationale de Sécurité du Médicament et des Produits de Santé (French health products safety agency)
App	Application
AQEL	App Quality Evaluation
BfArM	Bundesinstitut für Arzneimittel und Medizinprodukte
COPD	Chronic obstructive pulmonary disease
CE	Compliance with European community regulatory requirements
CEEBIT	Continuous Evaluation of Evolving Behavioural Intervention Technologies
CTTI	Clinical Trials Transformation Initiative
CV	Curriculum vitae app
DAP	Digital Assessment Portal
DAQs	Digital Assessment Questions
DGCCRF	Direction générale de la concurrence, de la consommation et de la répression des fraudes (French Directorate-General for Competition, Consumer Affairs and Prevention of Fraud)
DiGA	Digital Health Applications
DNS	Délégation ministérielle au numérique en santé (Ministerial delegation for digital healthcare)
DTAC	Digital Technical Assessment Criteria
ENS	Espace numérique de santé (Digital health space)
EQUATOR	Enhancing the QUality and Transparency Of health Research
EUnetHTA	European Network for Health Technology Assessment
FDA	Food and Drug Administration
GDPR	General Data Protection Regulation
GSMA	Global System for Mobile Communications
HAS	*Haute Autorité de santé* (French National Authority for Health)
HON	Health On the Net
HTA	Health Technology Assessment
JAMA	Journal of the American Medical Association
JMIR	Journal of Medical Internet Research
MARS	Mobile Application Rating Score
MD	Medical Device
MDM	Minimally disruptive care
mERA	mHealth Evidence Reporting and Assessment

MHAD	Mobile Health App Database
mHealth	Mobile Health
MMA	Mobile Medical Application
NCSC-FI	National Cyber Security Centre Finland
NHS	National Health Service
NICE	The National Institute for Health and Care Excellence
OBR	ORCHA Baseline Review
OMH	Our Mobile Health
ORCHA	Organisation for the Review of Care and Health Applications
PROMs	Patient-reported outcome measures
PREMs	Patient-reported experience measures
PwC	PricewaterhouseCoopers
RCP	Royal College of Physicians
RCT	Randomised controlled trial
RPM	Remote patient monitoring
UK	United Kingdom
WHO	World Health Organization
Wi-Fi	Wireless fidelity

Find all our publications at

www.has-sante.fr

HAUTE AUTORITÉ DE SANTÉ

Advancing quality in the fields of
health and social care services