Report of the Working Group on mHealth Assessment Guidelines

February 2016 – March 2017

1 INTRODUCTION

Following a call for expression of interest, the European Commission established a Working Group on mHealth assessment guidelines in February 2016. The mandate of the group was "to develop guidelines for assessing the validity and reliability of the data that health apps collect and process". Members of the Working Group, that were representatives from a broad range of stakeholders, were actively engaged in discussions and work preparing the production of the guidelines.

There was substantial knowledge exchange and sharing of relevant material, including on existing guidelines. This exercise was a useful learning process for all. Notably, it was valuable for getting an overview of the recent developments in the area of mHealth in Member States as well as mapping wide range of concerns of different stakeholders as regards validation of mHealth apps in general.

Most of the members were in favour of extending the scope for the guidelines i.e. assessing mHealth apps broadly and not only validity and reliability of their data. However, the Working Group did not come to an agreement on the way forward and therefore did not issue and endorse guidelines within its mandate.

2 SUMMARY OF THE PROCESS

With the support of Consard Limited, as facilitator and in charge of the drafting, the process was carried out involving 4 face-to-face meetings with the Working Group and 2 open stakeholder meetings, as well as conference calls and online questionnaires. In addition, the Working Group was supported by an online collaboration platform.

The overall timeline for the work foreseen was one year.

Working Group members and the European Commission provided feed-back after each new draft of the guidelines.

2.1 Working Group Activities

Phase 1

The Working Group met in Brussels for the first time on the 8th of March 2016 and again via a conference call on the 7th of April 2016. In preparation for the first meeting, members were asked to complete an online questionnaire to establish a basis for discussion on purpose, scope, and target groups for the guidelines.

A public consultation was conducted by means of an open online questionnaire.

Phase 2

The Working Group met in Brussels on the 3rd of May 2016 and again via a conference call on the 20th of May 2016.

An open stakeholder event, an open online questionnaire and the 2016 eHealth Week in Amsterdam were used for another open consultation.

The importance of case studies and alignment of the emerging guidelines with approaches being developed and adopted across European Union was agreed.

Phase 3

The Working Group held conference calls on the 16th of June and 7th of July 2016 and had a meeting in Brussels on the 14th of September. In the light of analysis of the responses to the public consultation received during phases 1 and 2, and based on the contents of draft guidelines, three sub-groups (SG) were formed to progress on three outstanding issues:

- **SG1**: Scope and Legal dovetailing
- **SG2**: Risk Assessment; Methodology and Scoring
- **SG3**: Targeting

Each SG had a leader and a deputy leader selected from the Working Group. Meetings were virtual and informal with no minutes. A progress report from each SG was presented at all cross-SG leaders coordination meetings. The outcome was a proposal on content to be included in the draft guidelines. Cross-SG leader coordination meetings were organised to secure consistency of the approach and to manage any cross-cutting issues.

Phase 4

A new draft of the guidelines was developed to take account of the work done by the SGs during the summer 2016. The Working Group then met in Brussels on the 10th of November. It was agreed that it would be essential at this final stage of the process to identify where a consensus existed within the WG on the areas to be covered. However, by the end of the meeting there was still not a firm agreement between stakeholder representatives, neither on the scope and target groups, nor whether the work should proceed on the basis of assessing both apps and data or only data as foreseen in the original mandate of the Working Group.

To address this question, Working Group members were requested to submit consolidated views from their respective constituency on relevant criteria to be used for assessment of validity and reliability (see annex 5.1). Substantial contributions were provided. Industry stakeholders primarily focused on providing assessment criteria to evaluate data validity and reliability, in line with the initial mandate of the Working Group as a starting point for the project, whereas all other stakeholders addressed a wider range of assessment criteria for apps.

The result of this exercise also revealed that stakeholders had not been able to reach a common understanding on key concepts. For example, stakeholders provided the following comments with regard to the meaning of the term "effectiveness" and whether this criterion should be included in the Guidelines:

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

- **Industry:** "It refers to the health outcomes and as such should be excluded from the guidelines". "It is likely that by the time an app effectiveness is fully assessed a new version of the app has been created, making the assessment outdated.
- **Public Authorities: "**A study of effectiveness requires a lot of time and during the intervening period the version of the app evaluated can become obsolete and could be a new improved version".
- **Payers and Social Care Insurance: "**Collecting fine data is not an end in itself. The app must be effective in fulfilling its purpose".
- **Research and Academia: "**A major barrier on the uptake of mHealth solutions within EU's healthcare systems, especially for healthcare providers, is the paucity of reliable evidence on the clinical benefits and cost-effectiveness of mHealth solutions. And even for fields of application where a number of studies exist, they are often not comparable (different study designs, inclusion criteria etc.) Studies on mHealth fail to provide persuasive results**"**.

Furthermore, no agreement between the stakeholder representatives was reached on this question of scope. This rendered the production and endorsement of guidelines within the mandate of the Working Group impossible.

Considering the responses from stakeholders (see section 3), the European Commission Unit CNECT H3 concluded that the required consensus on scope existed only in relation to assessment of data (i.e. there was no consensus to widen the scope). After a last consultation of the Working Group members following a conference call held on 10 January 2017, it was concluded that the gap between those who prefer guidelines on the assessment of mHealth app data and those who were in favour of assessing the app as a whole was too wide to pursue the exercise with the initial configuration.

2.2 Stakeholder consultations

The first online questionnaire was used to get a quick overview of the Working Group member's insights and opinions related to mHealth apps.

The first draft of the guidelines was then subject to consultation on three occasions:

- An open stakeholder meeting in Brussels on the 4th of May 2016.
- The EU eHealth stakeholder group meeting in Brussels on the 18th of May 2016.
- An open online questionnaire for all stakeholders ending on the 16th of May 2016.

The second draft of the guidelines was subject to consultation on three occasions:

- The eHealth Network meeting in Amsterdam on the 7th of June 2016.
- An open stakeholder meeting in Amsterdam on the 10th of June 2016.
- The eHealth Forum in Gastein (EHFG) 2016 in Austria on the 28th of September where the WG work was presented.

The open consultation for the third draft of the guidelines in December 2016 was cancelled because no consensus was found within the Working Group about the scope of the guidelines. Therefore, the guidelines were not ready for publication or open consultation.

3 STAKEHOLDERS' POSITIONS ON ASSESSMENT CRITERIA

Following the initial interest in broadening the scope of the original mandate, Members of the Working Group were invited to give their views on 13 selected assessment criteria:

1. Privacy
2. Transparency
3. Safety
4. Reliability
5. Validity
6. Interoperability
7. Technical stability
8. Effectiveness
9. Efficacy
10. Efficiency
11. Accessibility
12. Usability/desirability
13. Scalability

The questions for each possible criterion were:

- What is your understanding of this criterion?
- Is this relevant for assessing validity/reliability? Yes/No - Please justify
- Relevant examples that justify your position
- Assessment questions/comments

Views grouped in 6 stakeholder constituencies:

1. Patients
2. Healthcare Professionals
3. Industry
4. Public Authorities
5. Payers and Social Health Insurance
6. Academia and Research

The Industry responses addressed the relevance of the proposed criteria for assessment of data, whereas other stakeholder groups replied on to the relevance of the proposed criteria for assessment of apps.

3.1 Relevant criteria for assessment of *apps*

From the responses given, it appears that, with the exception of the industry representatives, that chose to limit its feedback to address data validity and reliability as initially foreseen by the mandate of the stakeholder group, a consensus view could be found on the relevance of these six criteria:

1. Privacy
2. Transparency
3. Reliability
4. Validity
5. Interoperability
6. Safety

Some stakeholders also agreed on the relevance of two other criteria for the assessment of apps:

- **Technical stability**: Patients, Public Authorities and Payers and Social Health Insurance.
- **Effectiveness**: Patients, Healthcare Professionals and Payers and Social Health Insurance (However there was a disparity in the understanding of this term).

Industry considered **Accessibility** as "worthy of consideration if the mandate was to provide an overall assessment of the app". Healthcare Professionals felt it was relevant. Payers and Social Health Insurance stated "all users of the target group of an app must be able to use it. In case there are limitations, the app should make them transparent, e.g. for persons with low vision." Industry considered **usability and desirability** as a market related issue, whereas the patient stakeholder responses proposed two additional criteria: **User experience, user centred design, suitability** and **Security**.

3.2 Some comments from stakeholders about specific criteria

This section illustrates the primarily challenges to agree on common terminology.

Transparency or Privacy? Issues relating to Transparency raised by many respondents are addressed by data protection law, specifically addressed by art 12-14 of the General Data Protection Regulation (GDPR). They may therefore be best considered under the Privacy criterion.

Transparency or Validity? Payers and Social Health Insurance raised the following issues under a discussion of the transparency criterion: "Where do the content and the scientific concept of the app come from?" "Where can the scientific documentation of the app and its algorithms be found?" These issues are directly relevant also to the assessment of data under the validity criterion.

Scalability or Interoperability? The following comment was made by Public Health Authorities "the most crucial aspect of scalability is being able to easily capture the results in the patient's EHR". This seems more relevant under the criterion interoperability.

Safety and Accuracy or Validity? Many of the concerns raised in stakeholder responses about safety, may in fact be better considered as the consequences of insufficient validity of data. Academia/Research and patient stakeholders raised the possibility of including accuracy as an additional criterion, as a sub-

criterion of validity or a separate criterion in its own right. In any case, it appeared that, accuracy certainly is an important aspect of validity and more specifically the accuracy of (data produced by) algorithms, is directly relevant to assessment of validity.

3.3 Relevant criteria for assessment of *data*

Only Industry responded on the relevance of the proposed criteria for assessment of data validity and reliability, linking the response to the original mandate of the WG.

Other stakeholder groups with the exception of Payers and Social Health Insurance, whose answers on the validity and reliability criteria did address data, did not express an opinion on relevant criteria for assessment of data. They based their responses on the broader question of relevance of criteria to the assessment of apps.

3.4 Basis for further work

The industry stakeholder response primarily addressed assessment criteria for data, whereas all other responses addressed assessment criteria for apps. It is difficult to draw a firm conclusion covering the perspectives of non-Industry stakeholders on the criteria for assessment of data (validity and reliability).

Referring to the responses made by stakeholder groups, other than Industry, on criteria for assessment of apps, it could be concluded that there is a consensus on the relevance of the following five criteria (assuming that there is an agreement on the meaning of these terms):

1. Privacy
2. Transparency
3. Reliability
4. Validity
5. Interoperability

Overall, there was a relatively wide variety of perspectives from stakeholders, areas of apparent disagreement and different understanding of the implications, use and meaning of the criteria during app assessment (see annex 5.1). In fact, the result of this exercise reveals that the work done during the initial phases of this process to agree on common terms and definitions was inconclusive.

It must be recalled that this process was designed to establish a shared view on scope, purpose and targets for voluntary assessment guidelines, as reflected in the early drafts of the guidelines. No agreement on extending the initial mandate of the Expert Group was achieved at that time.

4 CONCLUSIONS

The Working Group put a considerable amount of work into the objective of issuing guidelines but it appeared that building these guidelines was a much more complex exercise than expected at the beginning of the process, and that the work to be done goes beyond the original mandate of the Working Group. A minimal level of consensus between the members of the Working Group was not reached. It was thus impossible to achieve and endorse any guidelines.

Moreover, it is important to underline that the guidelines were to address a fast moving and evolving range of technologies. The potential, and some would argue disruptive, impact of the adoption of mHealth solutions on the healthcare and social care sectors has become clearer during the duration of the process. The policy responses to these developments by key stakeholders have also evolved. At European level, there were developments in some EU regulatory areas impacting mHealth apps: the new Medical Devices Regulation[1] and the new General Data Protection Regulation[2]. Also at Member State level, for instance in France and the UK guidelines and regulatory initiatives have evolved significantly during the same period.

Clearly, an important lesson from this exercise is the need to follow a step-wise approach, starting with a solid agreement on scope and terminology, especially if the Guidelines are to be developed by a multi-stakeholder group.

The experience of different stakeholders collected and documented by the Working Group is a usefulbasis for the future actions, notably by the eHealth Network.

[1] http://ec.europa.eu/growth/sectors/medical-devices/regulatory-framework/revision_en
[2] http://ec.europa.eu/justice/data-protection/reform/index_en.htm

ANNEXES

1. **Stakeholders' positions on assessment criteria**
2. **Some examples of case studies**

- Andalusia
- Catalonia
- Medappcare

3. **Some examples of existing guidelines**

- Our Mobile Health
- Andalusia: Complete list of recommendations on design, use and assessment of health Apps
- Catalonia: mHealth.cat
- France: Good practice guidelines on health apps and smart devices (mobile health or mHealth)
- Germany: Health apps & co: safe digital care products with clearer regulations
- United Kingdom: Guidance: Medical device stand-alone software including apps