

# Quality assessment and user experience analysis of digital health apps

Maciej Hyzy  
School of Computing,  
Ulster University,  
Newtownabbey  
BT37 0QB, UK  
hyzy-m1@ulster.ac.uk

Raymond Bond  
School of Computing,  
Ulster University,  
Newtownabbey  
BT37 0QB, UK  
rb.bond@ulster.ac.uk

Maurice Mulvenna  
School of Computing, Ulster  
University, Newtownabbey  
BT37 0QB, UK  
md.mulvenna@ulster.ac.uk

Lu Bai  
School of Computing,  
Ulster University,  
Newtownabbey  
BT37 0QB, UK  
l.bai@ulster.ac.uk

Simon Leigh  
ORCHA, Sci-Tech  
Daresbury, Vanguard  
House, Keckwick Lane,  
Daresbury WA4 4AB  
simon.leigh@orcha.co.uk

**This paper provides an overview of research conducted in the area of digital health app (DHA) evaluation focusing on user experience (UX) assessment. A literature review of existing evaluation frameworks has been conducted. Furthermore, in collaboration with ‘The Organisation for the Review of Care and Health Applications’ (ORCHA - a digital health app compliance company), a unique dataset comprising of quality assessment data for 2053 DHAs has been acquired. Using available literature and data science techniques on the ORCHA dataset, this research aims to derive new insights into the quality assessment process as well as new mobile UX knowledge and UX assessment tools.**

*mHealth. mHealth apps evaluation. Digital health. ORCHA assessment. UX assessment.*

## 1. INTRODUCTION

The amount of digital health apps (DHAs) has surpassed 350,000 [1] and the public interest in DHAs has been increasing [2]. For these apps to be safe to use by members of the public, they need to be quality assured. A common method used for evaluating DHAs are evaluation frameworks and the current literature suggests that improvements could be made to these frameworks [3]–[5]. Evaluation frameworks can be high level in nature and can be used to evaluate any DHA within any healthcare domain [5]. Evaluation frameworks use several different techniques such as Likert scales, calculations and checklists to evaluate DHAs [3].

## 2. RESEARCH

This research investigates the assessment process of DHAs including the evaluation frameworks for DHAs. Furthermore, ‘The Organisation for the Review of Care and Health Applications’ (ORCHA) has provided a dataset of 2053 DHAs that have been evaluated using their ORCHA baseline review version 6 (OBR V6) assessment tool [6]. This data is being used to support this PhD research. The OBR V6 assessment tool consists of approximately 350 questions, and based on these questions the app’s professional/clinical assurance, UX and data

privacy are assessed and accompanied with scores (including an overall ORCHA score). Data science methods will be used on this dataset together with the available literature with the research aim of learning about DHA assessment and providing data informed knowledge and tools that could be incorporated into the ORCHA assessment tool for enhancing the UX evaluation of DHAs.

### 2.1 Research aim

The aim of this research is to provide data informed ‘UX’ knowledge and tools to assess digital health apps.

### 2.2 Working research questions

1. What useful and actionable insights can be gained when analysing a large quality assessment dataset (ORCHA)?
2. Can we improve the user experience (UX) assessment process for digital health apps?
3. What is the inter-rater agreement between different professionals and end-users when assessing digital health apps (DHAs)?
4. Does UX experience of the assessor effect how the assessor scores the user interface (UI) /UX when using a checklist?

5. Can the assessment process of the digital health apps be automated? Or partially automated?

### 2.3 Evaluation frameworks for DHAs

Evaluation frameworks can be used to assess UX of DHAs. Evaluation frameworks provide an efficient method for assessing DHAs and can show where a DHA can be improved. Most of the evaluation frameworks consist of Likert scales (with different point scales) and dichotomous (polar) questions. Calculations are commonly computed to provide a quality score [3]. The data collected with evaluation frameworks lets us see the 'big picture' as all the data collected using the same evaluation framework allows for creation of datasets (such as the ORCHA dataset with 2053 health apps evaluated) and conducting descriptive statistics on such dataset to learn more about DHAs. Some of the evaluation frameworks such as the National Institute for Health and Care Excellence Evidence standards framework (NICE ESF) [7] set standards that DHAs should follow, and assign tiers (1, 2, 3a, 3b) based on app functionality and regulation.

### 2.4 Potential contribution

A possible contribution as a result of this research could include benchmarking of DHA performance, probability distributions in terms of professional & clinical assurance, data privacy and ORCHA scores and an improved UX evaluation for DHAs when using a checklist. Knowledge of inter-rater reliability between different professionals and end-users when assessing DHAs. Knowledge of whether a UX assessor's experience in the evaluation can affect their scoring of UI/UX when using a checklist will also be insightful.

### 2.5 Potential impact on stakeholders

ORCHA's DHA assessment tool (OBR V6) could be modified. DHA developers could be impacted as modified OBR V6 could be stricter in assessment of DHAs. The recommendations of DHAs to the public could change.

### 2.6 Limitations to the work

Machine learning (ML) techniques used to explore the dataset must be explainable (not black box). This is because lack of interpretability in ML models can undermine trust in DHAs.

### 3. Acknowledgements

This research is done in partnership with ORCHA, a UK-based digital health compliance company. This work is supported by a Northern Ireland DfE CAST award / PhD scholarship.

### 4. REFERENCES

- [1] "Digital health apps balloon to more than 350,000 available on the market, according to IQVIA report | MobiHealthNews."  
<https://www.mobihealthnews.com/news/digital-health-apps-balloon-more-350000-available-market-according-iqvia-report> (accessed Mar. 18, 2022).
- [2] S. Leigh et al., "Web-based internet searches for digital health products in the United Kingdom before and during the COVID-19 pandemic: a time-series analysis using app libraries from the Organisation for the Review of Care and Health Applications (ORCHA)," *BMJ Open*, vol. 11, no. 10, p. e053891, Oct. 2021, doi: 10.1136/BMJOPEN-2021-053891.
- [3] M. Hensher et al., "Scoping review: Development and assessment of evaluation frameworks of mobile health apps for recommendations to consumers," *Journal of the American Medical Informatics Association*, vol. 28, no. 6, pp. 1318–1329, Jun. 2021, doi: 10.1093/JAMIA/OCAB041.
- [4] P. Henson, G. David, K. Albright, and J. Torous, "Deriving a practical framework for the evaluation of health apps," *The Lancet Digital Health*, vol. 1, no. 2, pp. e52–e54, Jun. 2019, doi: 10.1016/S2589-7500(19)30013-5/ATTACHMENT/A19A88BB-16A5-40CC-B266-5595ED726F5C/MMC1.PDF.
- [5] S. Lagan, L. Sandler, and J. Torous, "Evaluating evaluation frameworks: a scoping review of frameworks for assessing health apps," *BMJ Open*, vol. 11, p. 47001, 2021, doi: 10.1136/bmjopen-2020-047001.
- [6] Hunt Sophie, "Review Documentation - Review Development & Resources | Exte." <https://confluence.external-share.com/content/b6055aac-83e4-4947-be0e-ebb8c39559ef> (accessed Mar. 13, 2022).
- [7] "Overview | Evidence standards framework for digital health technologies | Guidance | NICE," 2018. <https://www.nice.org.uk/corporate/ecd7> (accessed Nov. 06, 2021).