

Validity and reliability assessment of the Persian version of the User Mobile Application Rating Scale

Rezvaneh Sheybak¹, Rahele Samouei², Ahmad Papi³, Rasool Nouri³

¹Department of Medical Library and Information Science, Isfahan University of Medical Sciences, Isfahan, Iran, ²Social Determinants of Health Research Center, Isfahan University of Medical Sciences, Isfahan, Iran, ³Health Information Technology Research Center, Isfahan University of Medical Sciences, Isfahan, Iran

Background: Smartphone health applications have gained popularity among users; however, the accuracy of the health information they provide often goes unchecked, potentially compromising user safety. This highlights the need for assessing mobile health applications. While several evaluation tools exist, including the User Mobile Application Rating Scale (uMARS), there is currently no available Persian version of uMARS. This study aims to validate the Persian version of uMARS. **Materials and Methods:** This cross-sectional study involved students from the management and medical information science, modern medical technologies, and rehabilitation science programs at Isfahan University of Medical Sciences. Participants were selected using cluster random sampling and nonrandom quota sampling, resulting in a total of 106 undergraduate, master's, and PhD students. Nonrandom quota sampling was then conducted among undergraduate, master's, and PhD students from the aforementioned schools. One hundred and six students were chosen using nonrandom quota sampling. First, this questionnaire was translated by two English-Persian language expert translators and backtranslated, and this was compared with the original uMARS. After assessing face and content validity, convergent and divergent validity were assessed. To determine reliability, the Cronbach's alpha method and test-retest was used. **Results:** Face, content, convergent, and divergent validity scale of uMARS was affirmed, and internal consistency for all subjective and objective quality dimensions with Cronbach's alpha of 0.929 was affirmed. The reliability of all subscale test-retest was also acceptable. Considering calculated indicators, the Persian version of uMARS in this study has a good reliability and validity and is expected to be useful for mobile health applications as a standard tool. **Conclusion:** uMARS Persian version is a reliable tool to assess the quality of mobile health applications from the user's perspective. This assessment framework offers a comprehensive evaluation of various aspects of mobile health applications, including their engagement, functionality, esthetics, information, and subjective quality.

Key words: MHealth, mobile applications, mobile health, Persian language, User Version of the Mobile Application Rating Scale, validating

How to cite this article: Sheybak R, Samouei R, Papi A, Nouri R. Validity and reliability assessment of the Persian version of the User Mobile Application Rating Scale. J Res Med Sci 2026;31:17.

INTRODUCTION

Today, mobile tools, including smartphones and tablets, have become an essential part of our lives,^[1] and the number of their users is significantly increasing. A report by the Global System for Mobile Communications (GSMA) institution titled "mobile economy" published in 2025, says that by the end of 2024, 5.8 billion people (nearly 71% of the world

population) have been the customers and users of mobile phones. Consequently, it is predicted that mobile penetration will increase to 6.5 billion up to 2030 which will encompass 76% of the world population^[2] With the global rise in mobile phone usage, access to mobile health applications appears to significantly enhance patients' involvement in their recovery and treatment processes.^[3] Moreover, statistics indicate that one-fifth of smartphone users have installed at least one health-related application on their

Access this article online

Quick Response Code:



Website:

<https://journals.lww.com/jrms>

DOI:

10.4103/jrms.jrms_131_25

This is an open access article distributed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 License (CC BY-NC-ND), where it is permissible to download and share the work provided it is properly cited. The work cannot be changed in any way or used commercially without permission from the journal.

For reprints contact: WKHLRPMedknow_reprints@wolterskluwer.com

Address for correspondence: Dr. Rasool Nouri, Health Information Technology Research Center, Isfahan University of Medical Sciences, Isfahan, Iran. E-mail: nouri@mng.mui.ac.ir

Submitted: 20-Feb-2025; **Revised:** 26-Dec-2025; **Accepted:** 18-Feb-2026; **Published:** 29-Jul-2026

devices,^[4] and the users use these applications to increase their health and health-related knowledge. Specifically, after the outbreak of the COVID-19 pandemic, interest in these applications increased. It is expected that the applications in this field will increase in the near future. However, despite the fast growth of these applications market, the accuracy of hygienic information in these applications is rarely investigated by the supervision authorities,^[5] which can endanger the safety and health of users.^[6,7] Hence, the existence of a validated tool for assessing these applications is needed. One of the known tools for assessing mobile health applications, which has been noted in recent years, is star rating. However, the star ratings provided in app stores are inherently subjective,^[8] and it does not consider the essential and important content like the quality of information, the upgradeable characteristics and features, used for a long time by the users, and other essential points.^[9] As such, it is not an appropriate tool for assessing mobile health applications. Consequently, more robust assessment tools are necessary. Now, different tools have been built to assess mobile health applications; one of these tools is the Mobile Application Rating Scale (MARS) that was developed for evaluating the quality of mobile applications, and their psychological characteristics show the reliability and validity of the mentioned tool.^[10] However, MARS has been created for the use of researchers, physicians, and specialists^[11] and to use it needs to be an expert or a professional in the field of mobile health technology and also the area of related health^[8] and most users of mobile health applications are not specialists, the importance of including end-user assessments has been emphasized^[12-15] the need to the existences of one tool for assessing these applications by nonprofessional users also is seen. In this regard, the User MARS (uMARS), which is a derivative of the main MARS, is a questionnaire that suggests solutions for this problem. This version is simpler than MARS and features enhanced legibility.^[16] At present, Persian is the official language of Iran, Tajikistan, and one of the two official languages of Afghanistan (along with Pashto).^[17] However, about 110 million Persian speakers are in the world and Persian is among 20 superior world languages and the Persian speakers are growing in number and mobile health applications in Persian language is constantly increasing,^[18] given the rising use of mobile health applications in Persian-speaking countries, having a valid tool available in Persian can significantly assist researchers, developers, and users in efficiently assessing the quality of these applications. A recent study aimed to validate uMARS in the Persian language.

MATERIALS AND METHODS

Study design

The recent study employed a cross-sectional design. The research involved the translation and backtranslation of

the uMARS tool, with subsequent evaluation of its validity and reliability, utilizing a sample of students from Isfahan University of Medical Sciences. It should be noted that a fitness-related mobile health application was employed to evaluate the reliability of the study.

Description of the User Mobile Application Rating Scale

The uMARS tool has deleted the need to be a specialist and gives a reliable tool.^[19,20] Users can determine the positive and negative quality characteristics of health applications being rated, and uMARS would distinguish the scope that needs to be expanded more.^[21] This tool includes three objective quality parts (engagement, functionality, esthetics, and information) (questions 1–16), subjective quality (questions 17–20), and perceived impact conceptions (questions 20–26).^[16] The objective quality rating includes four dimensions: engagement (5 items), functionality (4 items), esthetics (3 items), information (4 items), and subjective quality rating 4 items. This tool was created in 2016 by Stoyanov *et al.*, assessing two applications informing the youth of alcohol and the effect of using music.^[16] In 2021, Martin-Payo *et al.* validated the Spanish version of this tool by assessing the application related to covid 19 prevention.^[19] Furthermore, Morselli *et al.* validated this tool in Italian by assessing Facebook in 2021.^[11] In 2022, the Japanese version of this tool was validated by Shinohara *et al.* in the application related to the function of pulmonary patients.^[9] In 2023, Chasiotis *et al.* validated the Greek version of uMARS by an application related to the health education of Thessaly University.^[20] Furthermore, using uMARS to assess application quality based on “prostate cancer,”^[22] “rheumatism patients,”^[23] and “dentistry”^[24] is shown.

Cross-cultural adaptation and translation process

To translate and backtranslate this tool, at first, the main English uMARS version was translated from English to Persian by two English expert translators, masters in information technology conceptions, who were Persian speakers. During this initial translation, there were no direct Persian equivalents for some certain terms; therefore, the closest meanings in Persian were utilized alongside the corresponding English terms, annotated as footnotes. Then the first created Persian version was backtranslated from Persian to English by two experts and was compared with the main English uMARS.

App assessment

First complete search in the Internet store named Cafe Bazar to find all mobile health applications relating to fitness was conducted. To identify the fitness applications, searching for the keyword fitness among all processed applications, the apps related to fitness were distinguished. Applications were categorized in the fields of sports, education, tools, books and press, medicine, entertainment, way of living, trading,

buying, and fitness. Up to this stage, 961 applications were identified (November 26th, 2022), and then, by entering the properties of each application, its characteristics were extracted. These characteristics include its name and its developer, the size of the application, the number of active installations, its price, the awards and privileges it has taken, the date of update, star justification, and the application classification. To identify the best application, they were screened according to the input criteria. The input criteria of the mentioned applications were the following: to be in the Persian language, related to fitness, to be free, receive at least 2.5 stars out of 5 stars via users, and to contain the best content among the applications with an installation number higher than 50,000 installations. Considering the criteria, 743 applications were omitted. The 218 remaining apps, according to the entrance and exit criteria, were classified. After application installation and examining the contents of the applications, the apps that solely had text or visual contents, related to one sex, did not have clear graphics, or were not validated originally were deleted. Twenty apps remained that according to the decision of the research team, the application Mancansho (to become thin in 30 days) was chosen as the selected application to enter the research; since it had all the questions and details first, it covered the four scopes of uMARS. Some of this application's characteristics are having suitable content related to fitness, both in text and visually, including a reminder, having source validation, calculating body mass index individually, and showing a weight diagram and calorie diagram for each person.

To affirm content validity, the questionnaire was given to 13 experts (4 health informatics, 5 health information management, 1 psychologist, and 3 physical education). The selection criteria for these specialists are based on their expertise, experience, and interdisciplinary diversity, all of which are essential for an effective content validity assessment of the uMARS questionnaire. The experts checked each subscale to assess uMARS content validity and assessed each of the items from the viewpoint of necessity, simplicity, clearance, and being relevant to the subject. The results of assessing 13 experts noted the calculated ratings, and according to the Lawshe formula, content validity ratio (CVR) and content validity index (CVI) were calculated for each item. The least acceptable amount of CVR according to the experts (13 persons), via the Lawshe technique for every scale item, is 0.54.^[25] According to this, the questions that the calculated number in the lawshe formula for them is bigger than 0.54 are distinguished and saved in necessary and important tools.

Formula 1: Calculating the CVR

$$CVR = \frac{ne - \frac{N}{2}}{\frac{N}{2}}$$

In this formula, "ne" indicates the number of persons who see the existence of the considered item necessary, and "N" shows all the skilled persons who have assessed the mentioned item.^[25]

Formula 2: Calculating the validity content indicator CVI

$$CVI = \frac{n}{N}$$

The amount CVI for each item <0.70 means not accepted, 0.70–0.79 means conditionally accepted, and more than 0.79 means the item is affirmed.^[26] Then, in the next process, the convergent and divergent validity was examined. For convergent validity, the correlation for each question is calculated with its related dimension, and for divergent validity, the correlation ratio of each question was calculated with each of the other three dimensions. In this section, the correlation amount of each question with its dimension should be more than the amount of correlation of every question with other questionnaire dimensions. Convergent validity is affirmed only if the correlation coefficient of one dimension is more than 0.40 with the dimension subscale itself, while more than 80% of correlation coefficients in each dimension subscale are higher than the correlation coefficients with other dimensions, then the divergent validity is affirmed.^[9] To assess the reliability of this scale, random cluster sampling for selecting the school of Isfahan Medical University was done, and the Management and Medical Information Science school, the modern medical technologies, and rehabilitation science were chosen, then nonrandom quota sampling among the bachelor, master and PhD students of the mentioned schools was done. Ninety-six people participated in the study.

The number of students from different schools was achieved by referring to the education department; then, the ratio of the number of students from different grades of each school was calculated, and all the obtained decimal numbers were rounded off upwardly, and at last, 98 students were considered as samples to enter the research [Table 1].

After suggesting to the students, 106 participants were intended to enter the research. The participants should be dominant Persian speakers, and they should intend to participate in the research. After selecting the participants the application called Mancansho (to become thin during 30 days) was given to them and the eligible persons were asked to check all the functions of the applications at least for 10 min and then using uMARS rate it, also to assess the reliability of this scale first 106 participants rated this application using uMARS and after 2 weeks 36 participants assessed the application again and responded to the uMARS and internal consistency of subscales and the total score of the subjective and objective quality was evaluated by using Cronbach's alpha. Cronbach's alpha of

Table 1: Population number and research sample of students from different schools

School	Grades	Number of students	Ratio of students
Management and medical information science	PhD	41	77.3
	Masters	75	89.6
	Bachelor	287	39.26
Rehabilitation sciences	PhD	17	56.1
	Masters	84	72.7
	Bachelor	435	40
New medical technologies	PhD	52	78.4
	Masters	53	87.4
	Bachelor	0	0

more than 0.70 is acceptable.^[27] To be confident of the reliability of the Persian uMARS assessment during 2 weeks' time, it was done twice that compared to the similar articles, this time delay was regarded as a suitable period of time by the research team. We did not examine 20–26 questions of uMARS that are related to the comprehended effect by the users in this study because they did not have the same nature, and as it is reported in the main uMARS, Japanese, and Greek studies.^[9,16,20] To analyze the data of this study, the SPSS version 26 (SPSS Inc., Chicago, IL, USA) was used.

RESULTS

One hundred and six people, including 95 women (89.6%) and 11 men (10.4%), participated in this research, of which 62.3% were bachelors, 28.3% masters, and 9.4% had PhD greats that 47.2% were from management and medical information science schools, 9.4% new medical technologies, and 43.4% rehabilitation sciences. After translating uMARS by translators, the translations were analyzed by the research team, and no major difference was found between the two translations. After discussion, the research team members resolved any minor differences, and to better and easily understand, the questionnaire was checked by nonexpert out of the research team, and the minor changes that this person suggested were made by the research team (Appendix 1; <https://links.lww.com/JRMS/A52>).

In terms of content validity, all items had acceptable CVR. Furthermore, considering the calculated points for CVI, most items were affirmed and only questions number 8, 9, 14 and 18 which from the simplicity the viewpoint got CVI = 0.76, and questions 9 and 12 that took CVI = 0.76 from the clearance aspect needed to be reviewed, also question 19 got CVI = 0.76 from the viewpoint of being relevant but considering that it had got enough CVR it remained in the Persian version of uMARS. In Table 2, the CVR and CVI for each of the uMARS questions are presented.

In addition, after reviewing the suggestions and comments from experts to revise and maintain the questions with a

Table 2: Content validity ratio and content validity index for each of the User Mobile Application Rating Scale questions

	CVR	CVI (simplicity)	CVI (clarity)	CVI (pertain)
1. Entertainment	0.69	0.92	1	0.92
2. Interest	0.69	0.92	0.84	0.92
3. Customization	0.84	1	1	1
4. Interactivity	1	0.84	0.92	1
5. Target group	0.84	0.84	0.84	0.92
6. Performance	0.84	1	0.92	0.92
7. Ease of use	1	1	1	1
8. Navigation	0.84	0.76	0.92	1
9. Gestural design	0.69	0.76	0.76	0.84
10. Layout	1	0.92	1	1
11. Graphics	1	0.84	0.92	0.92
12. Visual appeal	0.84	0.84	0.76	0.92
13. Quality of information	0.84	0.84	0.84	0.92
14. Quantity of information	0.69	0.76	0.84	1
15. Visual information	0.69	0.92	0.92	0.92
16. Credibility of source	0.84	0.84	0.92	0.84
17. Recommend this app to people	0.69	1	1	0.84
18. How many times do you think you would use this app	0.69	0.76	0.92	0.84
19. Pay for this app	0.69	0.92	1	0.76
20. Overall (star) rating of the app	0.69	0.92	0.84	0.92

CVR=Content validity ratio; CVI=Content validity index

CVI between 0.70 and 0.79, efforts were made to make the questions smoother and more readable with the necessary changes. For example, wherever the slash (/) appeared in the English version of uMARS, it was replaced in the Persian edition with appropriate words such as 'and' or 'or'. In the phase of convergent and divergent validity, all questions had the most correlation with their dimension ($P < 0.0001$) and the correlation was less with other dimensions, so the convergent and divergent validity was affirmed. Each of the components of the engagement dimension, including entertainment (0.76, 0.0001), content appeal (0.74, 0.0001), customization (0.84, 0.0001), interactivity (0.82, 0.0001), and target group (0.76, 0.0001), has a higher positive correlation with its dimension and is significant at the 0.0001 level.

Each of the components of the performance dimension, including performance (0.81, 0.0001), ease of use (0.66, 0.0001), navigation (0.81, 0.0001), and gestural design (0.76, 0.0001), has a higher positive correlation with its dimension and is significant at the 0.0001 level.

Each of the components of the esthetic dimension, including layout (0.79, 0.0001), graphics (0.85, 0.0001), and visual appeal (0.76, 0.0001), has a higher positive correlation with its dimension and is significant at the 0.0001 level.

Table 3: Calculating internal and external correlation dimensions in the User Mobile Application Rating Scale

Dimensions	Cronbach's alpha of dimensions	Cronbach's alpha if item deleted	Corrected item-total correlation	Correlation coefficient	ICC (CI)
Engagement	0.84	0.91	0.73	0.661	0.80 (0.58–0.88)
Functionality	0.76	0.86	0.82	0.652	0.74 (0.43–0.84)
Esthetics	0.72	0.87	0.86	0.701	0.70 (0.6–0.78)
Information	0.78	0.84	0.85	0.663	0.74 (0.7–0.78)
subjective quality	0.73	0.85	0.78	0.797	0.70 (0.59–0.79)
Cronbach's alpha (total)	0.929	-	-	0.793	0.89 (0.8–0.94)

ICC=Intraclass correlation coefficient; CI=Confidence interval

Each of the components of the information dimension, including information quality (0.78, 0.0001), information quantity (0.79, 0.0001), visual information (0.81, 0.0001), and source credibility (0.72, 0.0001), has a higher positive correlation with its dimension and is significant at the 0.0001 level.

Each of the components of the subjective quality dimension, including program introduction (0.70, 0.0001), usage (0.81, 0.0001), cost (0.69, 0.0001), and star rating (0.81, 0.0001), has a higher positive correlation with its dimension and is significant at the 0.0001 level.

To determine the internal consistency of uMARS, all participants, after installing and checking the chosen application, rated the application using uMARS. To determine test–retest reliability, 36 participants responded to the Persian version of uMARS again after 2 weeks, and all the obtained data were analyzed. No lost amount was among the answers. Table 3 shows the internal and external consistency of the uMARS dimensions.

The result shows that the internal consistency was acceptable for all dimensions ($\alpha = 0.72\text{--}0.84$) and excellent for the total score ($\alpha = 0.929$). The entire uMARS also demonstrated high reliability ($\alpha = 0.92\text{--}0.93$). In the test–retest reliability examination between every dimension during 2 weeks of performance, a strong correlation was calculated at $P < 0.0001$ level. Furthermore, the intraclass correlation coefficient was 0.893 (95% confidence interval: 0.825–0.940), $P < 0.0001$.

DISCUSSION

The present study was done to validate the Persian version of uMARS (Appendix). During the phase of transition and backtranslation, we did the same translation process for Mars Persian version,^[28] also by checking similar uMARS articles in Greek, Japanese, Italian, and Spanish languages; no major difference in the translation and backtranslation method was seen. After translating and backtranslating uMARS in Persian, since validity and reliability are the most important aspects of quantitative research and their exact investigation, it is known as a good and effective tool.^[29] Persian uMARS internal and external correlation scale, convergent and divergent, and content and face validity were

investigated. The findings showed that the uMARS content validity in 20 questions related to the subscale subjective and objective quality was affirmed by the experts, and the questions were recognized as necessary and remained in the questionnaire. Since content validity, to become sure that the questionnaire would examine what it is intended for choose the most important and correct content, and design a tool to the best shape by experts is analyzed.^[30] This result shows that the questions and dimensions of the questionnaire were presented clearly and were understandable and have a logical relationship together, and they examined one subject. In the stage of face validity affirmation, experts, after analyzing questionnaire questions and to understanding to questions better, presented their ideas and after checking by the research team the suggestions were applied. For example: to the expert's suggestion, the information in options b, g and d question 19 was filled and the question text: "do you pay for this application?" to "do you pay for this application? (Show your agreement from 1 (the least) to 5 (the most))" was changed and the options were rated like Likert scale, so since face validity affirmation specifies that questionnaire and its phrase are in relation with the aim of the study;^[31] so, Persian uMARS is in relation with the aim of the study. During the convergent validity determination, every question was assessed with its dimension, and the obtained result was that all questions had the most correlation with their dimension ($P < 0.0001$), also Japanese researchers resulted while validating Japanese uMARS that except questions 5 and 18 the convergent validity was acceptable for all subscale.^[9] Considering the meaning of convergent validity, if the correlation between the marks of tests that measure similar characteristics is high, the tool has convergent validity; the obtained results mean that Persian uMARS questions and dimensions show good correlation and similarity and examine one concept consistently, and there is a meaningful and positive correlation between questions with its dimension. Since divergent validity examines the correlation of one dimension with its nonrelevant dimensions,^[32] every question was examined with other dimensions in this study and the obtained result was that questions correlation extent of each dimension was less with other dimensions. In Japanese uMARS, the divergent validity only information subscale was affirmed.^[9] Hence, Persian uMARS convergent and divergent validity

is approved, and all Persian uMARS dimensions examine one consistent subject and do not test irrelevant variables. In addition, Kohzadi *et al.* conducted a study titled “Persian adaptation and validation of the uMARS,” where they examined the validation of the Persian version of uMARS for a meditation-related application. The results of this study also confirmed the adequate reliability and validity of this tool in assessing the quality of meditation applications.^[33] In reviewing the study conducted by Mohammadbeigi *et al.*, it is observed that the assessment of both convergent and divergent validity was not addressed in the validation of the Persian version of the uMARS tool. This omission represents a considerable limitation in the comprehensive evaluation of the tool’s validity. In contrast, the current study employs a more holistic approach to the validation of the uMARS tool by specifically examining convergent and divergent validity. This methodology enables researchers to achieve a deeper understanding of the tool’s efficacy and accuracy, thereby facilitating a more accurate interpretation of the results obtained. To determine reliability since the uMARS is a Likert scale and covers examining attitude, the best method to analyze reliability is using Cronbach’s alpha.^[31] The result of this study showed internal compatibility for all subjective and objective quality subscales with Cronbach’s alpha of 0.929. Furthermore, Japanese researchers reached the extent of 0.75–0.85 the penetration of Cronbach’s alpha in all subjective and objective qualities while validating Japanese uMARS.^[9] Furthermore, Cronbach’s alpha extent in similar studies such as original uMARS,^[16] Spanish uMARS,^[19] Italian uMARS,^[11] and Greek uMARS^[20] was acceptable. Considering Cronbach’s alpha shows how much it expresses alike concept scale and it has internal relation and scale dimensions and details,^[27] so affirming Cronbach’s alpha in this present study means that consisting details of Persian uMARS has good internal compatibility. At the final test–retest reliability affirmation phase, the result of this study showed the total mark of correlation coefficient of Persian uMARS was 0.793, also the test–retest reliability of main uMARS in the 1–2-month period was 0.66 and 0.70 for 3 months.^[16] In the Italian uMARS test–retest reliability, each question was analyzed with a Pearson correlation coefficient that showed high reliability with a significant $P = 0.05$ and a Pearson factor higher than 0.7.^[11] Noting that test–retest reliability measures the rate of being alike of the results during a given time and under similar work and conditions, Persian uMARS is acceptable from the consistency of the viewpoint, and this questionnaire is reliable for gathering data under similar conditions. It is expected that this Persian version would be useful for assessing mobile health applications by nonexpert users. Moreover, it will be useful for health experts because, with the combination with quality examination that they can examine with Mars, it allows them to advise their Persian mobile health applications with high quality.^[19] Therefore,

MARS and uMARS are available in the Persian language, and with the help of these two tools, experts and nonexpert users can assess mobile health applications. Furthermore, this study provides a suitable opportunity for similar studies in Persian-speaking countries in the world and can help to the expansion and increased knowledge in the field of mobile health.

Limitations

The Mancansho app was available only in an Android version, which prevented some students with Apple devices from participating in the study. This limitation is noteworthy in the context of our research. Moreover, all participants in this study were students and young individuals, and this age distribution does not reflect the general population. In addition, the study’s sample was predominantly female, creating a potential gender imbalance. To better generalize this scale, several accompanying health programs targeting different populations should be evaluated with this Persian version of the uMARS.

CONCLUSION

The results show that the uMARS Persian version is a reliable and credible tool for assessing mobile health applications and has acceptable validation and reliability. It is expected that Persian uMARS, along with Persian MARS, will be changed into a standard tool for assessing the quality of mobile health applications.

Financial support and sponsorship

This study was financially supported by the Isfahan University of Medical Sciences.

Conflicts of interest

There are no conflicts of interest.

REFERENCES

1. Rezaee R, Ghaffari M, Rabiei R, Kavousi A, Rakhshanderou S. Design and usability evaluation of a mobile application for self-care among Iranian adolescents. *BMC Public Health* 2024;24:892.
2. GSMA. The Mobile Economy; 2024. Available from: <https://www.gsma.com/solutions-and-impact/connectivity-for-good/mobile-economy/wp-content/uploads/2024/02/260224-The-Mobile-Economy-2024-infographic.pdf>. [Last accessed on 2025 Dec 19].
3. Diamantidis CJ, Becker S. Health information technology (IT) to improve the care of patients with chronic kidney disease (CKD). *BMC Nephrol* 2014;15:7.
4. Research2Guidance. mHealth App Developer Economics 2014. Berlin: Research2Guidance; 2014.
5. Iribarren SJ, Akande TO, Kamp KJ, Barry D, Kader YG, Suelzer E. Effectiveness of mobile apps to promote health and manage disease: Systematic review and meta-analysis of randomized controlled trials. *JMIR Mhealth Uhealth* 2021;9:e21563.
6. Najm A, Nikiphorou E, Kostine M, Richez C, Pauling JD, Finckh A, *et al.* EULAR points to consider for the development, evaluation

- and implementation of mobile health applications aiding self-management in people living with rheumatic and musculoskeletal diseases. *RMD Open* 2019;5:e001014.
7. Dantas LO, Carvalho C, Prando BC, McAlindon TE, da Silva Serrão PR. Mobile health technologies for the management of rheumatic diseases: A systematic review of online stores in Brazil. *Clin Rheumatol* 2021;40:2601-9.
8. 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:e27.
9. Shinohara Y, Yamamoto K, Ito M, Sakata M, Koizumi S, Hashisako M, *et al.* Development and validation of the Japanese version of the uMARS (user version of the mobile app rating system). *Int J Med Inform* 2022;165:104809.
10. 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.
11. Morselli S, Sebastianelli A, Domnich A, Bucchi C, Spatafora P, Liaci A, *et al.* Translation and validation of the Italian version of the user version of the mobile application rating scale (uMARS). *J Prev Med Hyg* 2021;62:E243-8.
12. Knitza J, Tascilar K, Messner EM, Meyer M, Vossen D, Pulla A, *et al.* German mobile apps in rheumatology: Review and analysis using the mobile application rating scale (MARS). *JMIR Mhealth Uhealth* 2019;7:e14991.
13. Knitza J, Simon D, Lambrecht A, Raab C, Tascilar K, Hagen M, *et al.* Mobile health usage, preferences, barriers, and ehealth literacy in rheumatology: Patient survey study. *JMIR Mhealth Uhealth* 2020;8:e19661.
14. Serlachius A, Schache K, Kieser A, Arroll B, Petrie K, Dalbeth N. Association between user engagement of a mobile health app for gout and improvements in self-care behaviors: Randomized controlled trial. *JMIR Mhealth Uhealth* 2019;7:e15021.
15. Baptista S, Oldenburg B, O'Neil A. Response to "development and validation of the user version of the mobile application rating scale (uMARS)". *JMIR Mhealth Uhealth* 2017;5:e16.
16. 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:e72.
17. Ahmadian BA. The Persian language and its functions in regional convergence in the Iranian world. *J Iran Stud* 2018;4:6-34.
18. Farsi Language History; 2023. Available from: <https://renaissance-translations.com/farsi-language-history/>. [Last accessed on 2025 Sep 17].
19. Martin-Payo R, Carrasco-Santos S, Cuesta M, Stoyan S, Gonzalez-Mendez X, Fernandez-Alvarez MD. Spanish adaptation and validation of the user version of the mobile application rating scale (uMARS). *J Am Med Inform Assoc* 2021;28:2681-6.
20. Chasiotis G, Stoyanov SR, Karatzas A, Gravas S. Greek validation of the user version of the mobile application rating scale (uMARS). *J Int Med Res* 2023;51:3000605231161213.
21. Ferguson MA, Maidment DW, Gomez R, Coulson N, Wharrad H. The feasibility of an m-health educational programme (m2Hear) to improve outcomes in first-time hearing aid users. *Int J Audiol* 2021;60:S30-41.
22. Adam A, Hellig JC, Perera M, Bolton D, Lawrentschuk N. Prostate cancer risk calculator mobile applications (Apps): A systematic review and scoring using the validated user version of the mobile application rating scale (uMARS). *World J Urol* 2018;36:565-73.
23. Lambrecht A, Vuillerme N, Raab C, Simon D, Messner EM, Hagen M, *et al.* Quality of a supporting mobile app for rheumatic patients: Patient-based assessment using the user version of the mobile application scale (uMARS). *Front Med (Lausanne)* 2021;8:715345.
24. Parvathy Thampy PS, Nijesh JE, Chaly PE, Shyam S, Vaishnavi S. Evaluation of mobile dental applications using user version mobile app rating scale (uMARS). *J Indian Dent Assoc Madras* 2020;7:1-6. [doi: 10.37841/jidam.2020.v7.i1.001].
25. Lawshe CH. A quantitative approach to content validity. *Pers Psychol* 1975;28:563-75.
26. Waltz CF, Bausell RB. *Nursing Research: Design, Statistics, and Computer Analysis*. Michigan: F.A. Davis Company; 1981.
27. Arabzoozani M, Hassani Pour S, Bayegi V. Understanding Cronbach's alpha: A necessity for implementation of original research studies. *Iran J Med Educ* 2014;14:831-2.
28. Shirvanian Dehkordi M, Samoui R, Zare Farashbandi F, Nouri R. Psychometrics of the Farsi Version of the Companion Health Application Rating Scale in the Field of Pregnancy. *The First International Telemedicine Conference*; 1400.
29. Masuwai A, Zulkifli H, Hamzah MI. Evaluation of content validity and face validity of secondary school Islamic education teacher self-assessment instrument. *Cogent Educ* 2024;11:2308410.
30. Shabankareh K, Hamidi A. Validity assessment in the knowledge & information science researches. *Iran J Inf Process Manage* 2019;34:1103-24.
31. Mohammadbeigi A, Mohammadsalehi N, Aligol M. Validity and reliability of the instruments and types of measurements in health applied researches. *J Rafsanjan Univ Med Sci* 2015;13:1153-70.
32. Salehi M, Zarei S, Jabarpour E. Investigating the effect of green customers' integration on green innovation performance considering the mediating role of company size and environmental awareness of senior managers the case of Bahman company. *Res Prod Oper Manage* 2024;15:51-69.
33. Kohzadi Z, Rahmatizadeh S, Kohzadi Z, Valizadeh-Haghi S. Persian adaptation and validation of the user version of the mobile application rating scale (uMARS). *PLoS One* 2025;20:e0320349.

APPENDIX

نسخه فارسی مقیاس رتبه‌بندی برنامه‌های کاربردی همراه ویژه کاربران (uMARS)

دستورالعمل استفاده:

ارزیابان باید:

- الف) از اپلیکیشن استفاده کنند و حداقل 10 دقیقه آن را به طور کامل آزمایش کنند؛
- ب) تعیین کنند که استفاده از آن چقدر آسان است، چقدر خوب عمل می‌کند و آیا در راستای هدفی که برای آن ساخته شده است، عمل می‌کند؟
- ج) تنظیمات اپلیکیشن، اطلاعات توسعه‌دهنده، پیوندهای خارجی، ویژگی‌های امنیتی و غیره را بررسی کنند.

امتیازدهی

الف: میانگین امتیاز درگیرسازی:

ب: میانگین امتیاز عملکرد:

ج: میانگین امتیاز زیبایی‌شناسی:

د: میانگین امتیاز اطلاعات:

* سؤالاتی که با عنوان «کاربرد ندارد (N/A)» ارزیابی شده‌اند را از محاسبه میانگین امتیاز حذف کنید.

میانگین امتیاز کیفیت اپلیکیشن (الف+ب+ج+د)/4=

بسته به اهداف پژوهش، مقیاس کیفیت ذهنی از اپلیکیشن را می‌توان به صورت موارد جداگانه یا به عنوان میانگین امتیاز گزارش کرد.

موارد تأثیر درک شده را می‌توان تعدیل کرد و برای به دست آوردن اطلاعاتی در مورد تأثیر درک شده اپلیکیشن بر دانش، نگرش و قصد کاربران مرتبط با رفتار سلامت مورد نظر استفاده کرد.

نام اپلیکیشن:

دور عددی که کیفیت اپلیکیشن ارزیابی شده توسط شما را دقیق‌تر بیان می‌کند، خط بکشید. همه موارد در مقیاس پنج‌درجه‌ای از 1، ناکافی» تا «5، عالی» رتبه بندی شده‌اند. اگر اجزایی از اپلیکیشن نامرتبط می‌باشد، گزینه «کاربرد ندارد» را انتخاب نمایید.

رتبه‌بندی کیفیت اپلیکیشن

بخش الف

درگیرسازی - سرگرمی، جذابیت محتوایی، سفارشی‌سازی، تعامل‌پذیری، دستورات (به عنوان مثال، ارسال هشدارها، پیام‌ها، یادآوری‌ها، بازخوردها، قابلیت اشتراک گذاری) را دارد.

1- سرگرمی: اپلیکیشن سرگرم‌کننده است؟ آیا اجزایی دارد که آن را نسبت به سایر اپلیکیشن‌های مشابه سرگرم‌کننده‌تر کند؟

الف) کسل‌کننده است. د) نسبتاً سرگرم‌کننده است.

ب) عمدتاً خسته‌کننده است. ه) بسیار سرگرم‌کننده است.

ج) تا اندازه‌ای سرگرم‌کننده و لذت‌بخش است.

2- جذابیت محتوایی: استفاده از اپلیکیشن جذابیت لازم را دارد؟ آیا اطلاعات خود را در مقایسه با سایر اپلیکیشن‌های مشابه به شکل جالبی ارائه می‌دهد؟

الف) اصلاً جذابیت ندارد. (د) نسبتاً جذاب است.

ب) عمدتاً جذابیت ندارد. (ه) بسیار جذاب است.

ج) خوب است.

3- سفارشی‌سازی: به شما این امکان را می‌دهد که تنظیمات و اولویت‌هایی را که می‌خواهید سفارشی کنید (برای مثال صدا، محتوا، اعلان‌ها)؟

الف) امکان هیچ‌گونه سفارشی‌سازی را نمی‌دهد. (د) گزینه‌های متعددی را برای سفارشی‌سازی فراهم می‌کند.

ب) امکان سفارشی‌سازی کمی را فراهم می‌کند. (ه) امکان تنظیم کامل را فراهم می‌کند.

ج) امکان سفارشی‌سازی پایه را ارائه می‌دهد.

4- تعامل‌پذیری: آیا امکان ورودی کاربر، ارائه بازخورد، دستورات (یادآورها، گزینه‌های اشتراک‌گذاری، اعلان‌ها و غیره) را برقرار می‌کند؟

الف) بدون ویژگی‌های تعاملی است. (د) انواع ویژگی‌های تعاملی را ارائه می‌دهد.

ب) برخی ویژگی‌های تعاملی را دارد. (ه) سطح بسیار بالایی از پاسخگویی فراهم می‌کند.

ج) ویژگی‌های اولیه تعاملی را دارد.

5- گروه هدف: آیا محتوای اپلیکیشن (دیداری، زبان، طراحی) برای مخاطبین هدف مناسب است؟

الف) کاملاً نامناسب، نامشخص یا گیج‌کننده است. (د) متناسب با نیاز مخاطب هدف طراحی شده است.

ب) عمدتاً نامناسب، نامشخص یا گیج‌کننده است. (ه) به طور خاص متناسب با نیاز مخاطب هدف طراحی شده است.

ج) قابل قبول است.

بخش ب

عملکرد - سهولت استفاده، هدایت‌گری، روند منطقی و طراحی ایمايي اپلیکیشن

6- عملکرد: ویژگی‌ها (عملکردها) و اجزا (دکمه‌ها/منوها) اپلیکیشن تا چه حد دقیق و سریع عمل می‌کنند؟

الف) اپلیکیشن اجرا نمی‌شود. (د) عمدتاً عملکردی قابل قبول دارد.

ب) برخی از عملکردها کار می‌کنند. (ه) پاسخ کامل و به موقع است.

ج) روی هم رفته اپلیکیشن کار می‌کند.

7- سهولت استفاده: یادگیری نحوه استفاده از اپلیکیشن تا چه حد آسان است؛ برچسب‌های منو، نمادها و دستورالعمل چقدر واضح هستند؟

الف) دستورالعملی ندارد یا در صورت داشتن محدود است. (ج) تا حدی زمان یا تلاش نیاز دارد.

ب) زمان یا تلاش زیادی نیاز دارد. (ه) بلافاصله می‌توان اپلیکیشن را استفاده کرد.

(د) یادگیری آن آسان است.

8- هدایت‌گری: آیا حرکت بین صفحه نمایش منطقی است؟ آیا اپلیکیشن همه پیوندهای (لینک‌های) لازم بین صفحات را دارد؟

(الف) اصلاً ارتباط منطقی بین صفحات وجود ندارد.

(د) برای تعیین مسیر، درک آن آسان است.

(ب) بعد از زمان و تلاش زیاد قابل‌درک است.

(ه) جریان صفحه نمایش کاملاً منطقی، آسان، واضح و قابل‌درک است.

(ج) بعد از مدتی زمان و تا حدی تلاش قابل‌درک است.

9- طراحی ایمایی اپلیکیشن: آیا ضربه زدن روی صفحه نمایش (taps)، کشیدن انگشت به سرعت روی صفحه نمایش (swipes)، تغییر اندازه یک شیء با کلیک همزمان بوسیله دو انگشت (pinches) و حرکت دادن متن یا تصویر نمایش داده شده در جهت خاصی روی صفحه نمایش (scrolls) عمل می‌کند؟ آیا در همه اجزاء، صفحات نمایش باثبات هستند؟

(الف) کاملاً بدون ثبات و گیج‌کننده است.

(د) با مشکلات جزئی عمدتاً باثبات و قابل‌درک است.

(ب) اغلب بدون ثبات و گیج‌کننده است.

(ه) کاملاً باثبات و قابل‌درک است.

(ج) با وجود برخی بی‌ثباتی‌ها یا عناصر گیج‌کننده، خوب است.

بخش ج

زیبایی‌شناسی - طراحی گرافیکی، جذابیت ظاهری، طرح رنگ و ثبات سبک

10- طرح بندی: چیدمان و اندازه دکمه‌ها، نمادها، منوها و محتوای روی صفحه مناسب است؟

(الف) طراحی بسیار بد و نامنظم است.

(د) اغلب واضح است.

(ب) طراحی بد، تصادفی و نامشخص است.

(ه) به‌صورت حرفه‌ای، سازماندهی شده است.

(ج) رضایت‌بخش است.

11- گرافیک: کیفیت و وضوح گرافیک مورد استفاده برای دکمه‌ها، نمادها، منوها و محتوا تا چه حد بالا است؟

(الف) گرافیک‌ها غیرحرفه‌ای به نظر می‌رسد.

(ب) گرافیک‌ها و طراحی دیداری دارای وضوح و کیفیت پایین است.

(ج) گرافیک‌ها و طراحی دیداری، کیفیت متوسطی دارد.

(د) گرافیک‌ها و طراحی دیداری دارای وضوح و کیفیت بالا است.

(ه) گرافیک‌ها و طراحی دیداری دارای وضوح و کیفیت بسیار بالا است.

12- جذابیت ظاهری: ظاهر اپلیکیشن چقدر خوب است؟ اپلیکیشن چه اندازه خوب به نظر می‌رسد؟

(الف) زشت (ب) بد (ج) خوب (د) گرافیک یکپارچه (ه) زیبا

اطلاعات - حاوی اطلاعات باکیفیت بالا (مانند متن، بازخورد، سنجه‌ها، منابع و مآخذ) از یک منبع معتبر است.

- 13- کیفیت اطلاعات:** آیا محتوای اپلیکیشن صحیح است، به خوبی نوشته شده و با هدف و موضوع اپلیکیشن مرتبط است؟
- الف) نامرتب، نامناسب، نامنسجم و نادرست است.
 ب) ضعیف، به‌سختی مرتبط، مناسب یا منسجم است.
 ج) نسبتاً مرتبط، مناسب، منسجم و درست است.
 د) مرتبط، مناسب، منسجم و درست است.
 هـ) بسیار مرتبط، مناسب، منسجم و درست است.

14- کمیت اطلاعات: آیا اطلاعات ارائه شده در اپلیکیشن جامع و در عین حال مختصر است؟

- الف) حجم اطلاعات بسیار کم یا بسیار زیاد است.
 ب) حجم اطلاعات کم یا احتمالاً زیاد است.
 ج) خوب است؛ اما جامع و در عین حال مختصر نیست.
 د) اطلاعات وسیعی پیشنهاد می‌کند.
 هـ) جامع و در عین حال مختصر است.

15- اطلاعات دیداری: آیا شرح دیداری مفاهیم (از طریق نمودارها، گراف‌ها، تصاویر، ویدئوها و غیره) واضح، منطقی و صحیح است؟

الف) کاملاً نامشخص
 ب) عمدتاً نامشخص
 ج) خوب
 د) اکثراً واضح
 هـ) کاملاً واضح

- 16- اعتبار منبع:** آیا به نظر می‌رسد که اطلاعات داخل اپلیکیشن از یک منبع معتبر باشد؟
- الف) کاملاً نامعتبر است.
 ب) اعتبار کمی دارد.
 ج) نامعتبر نیست، اما معتبر بودن منبع هم واضح نیست.
 د) احتمالاً از یک منبع معتبر و قانونی آورده شده است.
 هـ) قطعاً از یک منبع قانونی و تخصصی آورده شده است.

کیفیت ذهنی اپلیکیشن

17- آیا این اپلیکیشن را به افرادی که ممکن است برای آن‌ها مفید باشد معرفی می‌کنید؟

- الف) اصلاً، این اپلیکیشن را به کسی توصیه نمی‌کنم.
 ب) این اپلیکیشن را به افراد بسیار کمی توصیه می‌کنم.
 ج) شاید، این اپلیکیشن را به برخی افراد توصیه کنم.
 د) این اپلیکیشن را به افرادی زیادی توصیه می‌کنم.
 هـ) قطعاً، این اپلیکیشن را به همه توصیه می‌کنم.

18- اگر اپلیکیشن پاسخگوی نیاز شما بوده، فکر می‌کنید چند بار در سال از آن استفاده کنید؟

- الف) اصلاً ب) 2 بار ج) 10 بار د) 10-50 بار هـ) بیش از 50 بار

19- آیا برای این اپلیکیشن هزینه‌ای پرداخت می‌کنید؟ (میزان موافقت خود را از 1 (کمترین) تا 5 (بیشترین) مشخص کنید).

- الف) 1 ب) 2 ج) 3 د) 4 هـ) 5

20- به طور کلی از پنج ستاره، چه امتیازی به این اپلیکیشن می‌دهید؟

- الف) * (خیلی ضعیف)
 ب) ** (ضعیف)
 ج) *** (متوسط)
 د) **** (خوب)
 هـ) ***** (عالی)

تأثیر درک شده از اپلیکیشن

بخش و

1- آگاهی - استفاده از این اپلیکیشن، آگاهی من را در مورد اهمیت بهبود رفتار سلامت افزایش داده است.

کاملاً مخالف	1	2	3	4	5	کاملاً موافق
--------------	---	---	---	---	---	--------------

2- دانش - این اپلیکیشن دانش و فهم مرا از رفتار سلامت افزایش داده است.

کاملاً مخالف	1	2	3	4	5	کاملاً موافق
--------------	---	---	---	---	---	--------------

3- نگرش - این اپلیکیشن نگرش مرا نسبت به بهبود رفتار سلامت تغییر داده است.

کاملاً مخالف	1	2	3	4	5	کاملاً موافق
--------------	---	---	---	---	---	--------------

4- قصد تغییر - این اپلیکیشن قصد و انگیزه مرا برای توجه به رفتار سلامت افزایش داده است.

کاملاً مخالف	1	2	3	4	5	کاملاً موافق
--------------	---	---	---	---	---	--------------

5- جستجوی کمک و استفاده از این اپلیکیشن، من را تشویق به جستجوی کمک بیشتر برای بهبود رفتار سلامت می کند (اگر به آن نیاز داشته باشم).

کاملاً مخالف	1	2	3	4	5	کاملاً موافق
--------------	---	---	---	---	---	--------------

6- تغییر رفتار - استفاده از این اپلیکیشن باعث افزایش یا کاهش رفتار سلامت می شود.

کاملاً مخالف	1	2	3	4	5	کاملاً موافق
--------------	---	---	---	---	---	--------------

ارائه نظرات بیشتر در مورد این اپلیکیشن