Opinion Paper

Snežana Jovičić*, Joanna Siodmiak and Ian D. Watson, on behalf of the European Federation of Clinical Chemistry and Laboratory Medicine Working Group on Patient Focused Laboratory Medicine

Quality evaluation of smartphone applications for laboratory medicine

<https://doi.org/10.1515/cclm-2018-0710>

Received July 6, 2018; accepted November 7, 2018; previously published online November 29, 2018

Abstract

Background: Many of the mobile applications (apps) used for delivering health interventions involve laboratory medicine data. This survey was conducted to search the online market for health apps that manage laboratory medicine data. The aim was to review them and perform a quality evaluation.

Methods: Apps search criteria were “Lab results blood work”, “Lab results”, and “Health apps”. After the step-wise exclusion process, 52 selected apps were downloaded and analyzed. For review and content analysis of the apps, a multidimensional tool for classifying and rating the quality of mobile health apps – Mobile App Rating Scale (MARS), was used.

Results: Selected apps were classified into five categories according to their intended use by patients or physicians, and the type of data engaged. Spearman’s correlation analysis found significant correlations between MARS individual scoring items, as with the subjective quality and number of technical aspects. Kruskal-Wallis analysis showed a significant difference in the number of technical aspects employed, MARS engagement and informational quality score items, total score, and subjective quality. The lowest values for all of these items were in the category of

apps designed for patients, and the post hoc test showed that the difference was statistically significant between this and the values in all other categories.

Conclusions: Apps designed for patients, are of the poorest quality, considering the total quality of the content and information they provide, estimated using the MARS tool. This estimation needs to be validated for laboratory medicine apps, and eventually modified after consideration of specific quality benchmarks.

Keywords: app; Mobile App Rating Scale (MARS); mobile health application; smartphone.

Introduction

By the end of the 20th century the interest of patients to participate in making decisions regarding their own health has increased [1]. This required a facilitation of access to information relevant to healthcare. Among the most important information are laboratory test results, as a significant number of clinical decisions are based on them [2]. The study performed by the European Federation of Clinical Chemistry and Laboratory Medicine Working Group on Patient Focused Laboratory Medicine showed that about 30% of patients wanting additional information on laboratory results would refer to the Internet [3].

On the other hand, the rapid increase of the global use of smartphones, thanks to their portability, facilitates access to health information at any time. Namely, technical capabilities of smartphones have been advancing rapidly, enabling many of the additional functions, beyond simple communication. The introduction of smartphones and their software platforms (Apple’s iOS, Google’s Android, Microsoft Windows Phone) were a starting point for designing special purpose applications, known as mobile applications (apps) [4]. Many of these apps are used for delivering health interventions and involve, in different ways, laboratory medicine data.

*Corresponding author: Snežana Jovičić, Center for Medical Biochemistry, Clinical Center of Serbia, Višegradska 26, 11000 Belgrade, Serbia, Phone/Fax: +381 11 361 56 31; and Department for Medical Biochemistry, Faculty of Pharmacy, University of Belgrade, Belgrade, Serbia, E-mail: snezanaj@pharmacy.bg.ac.rs, hionati@gmail.com

Joanna Siodmiak: Department for Laboratory Medicine, Faculty of Pharmacy, Collegium Medicum in Bydgoszcz, Nicolaus Copernicus University in Torun, Bydgoszcz, Poland

Ian D. Watson: c/o European Federation of Clinical Chemistry and Laboratory Medicine, Milan, Italy

The increasing number of available sources of health care information, either through web sites or mobile apps, triggers the need of the appraisal of their quality. Several quality assessment methods are available for health-related web sites and Internet information [5, 6]. Among them are the guidelines of the National Institute of Health's National Library of Medicine, which include the disclosure of site provider and funding, the evaluation of quality of the content through identifying the source of information and the expert involvement in the review of the content, as well as the declared privacy policy for manipulation with any of the users' personal information [7].

The quality of health-related apps, on the other hand, is more challenging. There have been attempts to apply tools developed for the evaluation of quality of web sites, as well as methods for systematic review of the literature, for identifying and evaluating health apps for effective health promotion and chronic disease management [4, 8–13].

Considering the importance of laboratory medicine data in the overall healthcare-decision making, this survey was conducted to search the online stores market for Android (Google Play) and iPhone (iTunes), as two of the most widely used smartphone platforms, for health apps that manage in any way user's laboratory medicine data. The aim was to review them and perform a quality evaluation.

Materials and methods

Study selection

We used the proposed method for the analysis of the content of mobile health apps by Grundy et al. [10], who recommended reporting standards by BinDhim et al. [11] and the use of multidimensional quality assessment tools, like the Mobile App Rating Scale (MARS) [12]. The apps for the analysis were collected throughout the period January 2017–May 2018, and the app search was conducted within Google Play and iTunes in Serbia. The search criteria entered in Google Play and App Store were “Lab results blood work”, “Lab results”, and “Health apps”. These criteria yielded 1247 potentially relevant apps, 747 for Android and 500 for iPhone, out of which 16 were available for both platforms. Apps were included if they, in any way, dealt with laboratory medicine data, had an English language interface and were available for smartphones. Out of 1247 potentially relevant apps, 95 were included in further analysis (29 for Android, 50 for iPhone, and 16 for

both), after the exclusion of the “prank apps” which allegedly measure the blood concentration of glucose or cholesterol through a finger scan, and apps that only mention laboratory data, but actually do not operate with them. In this phase, the information about apps was collected from their description in the store. In order to review every selected app, the apps were downloaded and analyzed based on their content and functioning. During this process we needed to further exclude apps that requested payment of a certain fee prior to download, apps that could not be started after the download and those that requested a specific login data. Finally, we were able to analyze the content and function of 52 selected apps – 27 for Android platform, nine for iPhone, and 16 available on both of them. The flowchart of the selection process is presented in Figure 1.

Review and content analysis of the apps

Descriptive data collected about each app were app name, version, developer, price, average ratings, total number of ratings, affiliation, category, and the year of the last update. Technical aspects of the app included the possibility of sharing on social media, whether it had an app community, does it allow password protection, require login, send reminders, and need web access to function. For the technical aspect, the total number of present aspects was calculated and used for further analysis. For review and content analysis of the apps we used a multi-dimensional tool for classifying and rating the quality of mobile health apps – MARS, developed by Stoyanov et al. [12]. The apps quality criteria are systematized into four categories divided into 19 subcategories representing the 19 individual MARS items: 1. Engagement (entertainment, interest, customization, interactivity, and target group), 2. Functionality (performance, ease of use, navigation, gestural design), 3. Aesthetics (layout, graphics, visual appeal), and 4. Information quality (accuracy of app description, goals, quality and quantity of information, visual information, credibility, evidence base). Each item uses a 5-point scale (1-Inadequate, 2-Poor, 3-Acceptable, 4-Good, 5-Excellent). If an item cannot be applicable for all apps, there is a “Not applicable” option. MARS is scored by calculating the mean score for each of the five categories. The fifth category describes app subjective quality, and it is scored as a mean subjective quality score out of personal raters' opinion on whether they would recommend the app, how many times they would use it, whether they would pay for it, and of the overall star rating of the app. The MARS scoring was conducted independently by

Figure 1: Flow diagram of the process of app selection for evaluation of function and content.

the two co-authors (SJ and JS), both laboratory medicine specialists. The internal consistency of MARS ratings was estimated with calculation of Cronbach's alpha. Interrater reliability was determined by calculating the intraclass correlation coefficient (ICC). Values of $ICC < 0.5$ indicate poor interrater reliability, values between 0.5 and 0.75 are considered moderate, 0.75–0.9 good, and > 0.9 indicated excellent agreement in MARS scoring. Calculation of ICCs was performed as two-way mixed effects, average measures model with absolute agreement [14, 15].

Calculated scores were then used for appropriate statistical analysis. We used Spearman's correlation analysis to analyze the relationships among user's ratings found in the app stores, quality elements defined by the MARS score, technical aspects of the app, as well as the affiliation of the developer. Also, the Kruskal-Wallis analysis of differences between the number of technical aspects present, MARS score items, and users' ratings of apps classified in the distinguished categories of apps using laboratory medicine data was performed. Statistical analysis was carried out using IBM® SPSS® Statistics (version 20) and MedCalc Software (version 18).

Additionally, we have added in the analysis the elements that we have considered particularly important for the field of laboratory medicine – comparability with locally accepted reference ranges (yes/no), units of concentrations used (SI/conventional units/both), the use of appropriate terminology (yes/no), and transfer security for uploading/downloading data, defined with clear and visible statement of privacy policy (yes/no). Also, we have included the affiliation of the app producer (classified as unknown, commercial, government, non-governmental organization – NGO, university), as an indicator of the credibility of the information used within the app.

Results

Apps categorization

All of the 95 eligible apps could be categorized into one of the seven categories according to their purpose described in the app store. These categories were:

1. apps that offer medical advice about symptoms and health queries with the possibility to upload laboratory test results, which can be seen, stored and shared (9/95, 9.5%),
2. reference ranges of selected analysis with basic information about the causes of increase or decrease designed for patients (15/95, 15.8%),
3. quick reference for laboratory tests for medical students and doctors (30/95, 31.6%),
4. apps for monitoring the state of user's health through a wide range of health parameters, including glucose and/or cholesterol as laboratory data (19/95, 20.0%),
5. apps that provide access to patients' laboratory results to physicians (11/95, 11.6%),
6. apps that enable patients to access their laboratory test results directly from the diagnostic center (4/95, 4.2%), and
7. electronic health records apps that include laboratory test results (7/95, 7.4%).

However, all of the items classified into the categories of apps that provided access to patient's results to physicians, and mobile electronic health record apps, after download allowed access only with specific login data, and therefore they could not be fully evaluated. Therefore, instead of seven, we could classify the analyzed apps into only five categories. Also, from the category of apps that allowed patients' access to their laboratory test results directly from the providing laboratory, only one app could be evaluated following the download. All of the included apps ($n = 52$) are listed in Table 1.

Summary characteristics of the selected apps

The descriptive data for the analyzed apps are presented in Table 2.

Most of the apps had none of the technical aspects examined (29/52), only one followed in frequency (10/52), then two (6/52), three (3/52), while four and five were present only in two apps each (2/52). None of the examined apps had all six technical features employed. Among less than half of the examined apps that did have one or more of the features of interest, most of them had enabled possibility of sharing data, and the least had an app community. MARS ratings performed by the two independent raters demonstrated high internal consistency (Cronbach's alpha for individual MARS categories ranged from 0.5 to 0.88, and for the total MARS score was 0.93). Interrater variability was moderate to good for individual

MARS categories (ICC between 0.5 and 0.83), as for the total MARS score (ICC 0.86). The average MARS total score was 3.8 out of 5 (IQR 0.8), ranging from 2.2 (Blood Tests Result by King of Story App) to 4.8 (Biolab). The highest individual MARS scoring item was for the Functionality 5.0 (IQR 0.0), and the lowest was Engagement, with median individual score value of 2.6 (IQR 1.2). Aesthetics and Information quality had considerably high and equal scoring values of 3.7 (IQR 0.7) and 3.8 (IQR 1.0), respectively. Interestingly, the Subjective quality score was lower than the total MARS score, with the value of 3.2 (IQR 1.0).

As it can be seen from the summary statistics presented in Table 2, most of the apps were using either conventional units or both SI and conventional units for the presentation of laboratory values. By both, we have considered either the possibility of choosing the type of units in which the user wants the values to be presented, or the presentation of some values in SI, and others in conventional units. Some of the apps of the second group offered the values of conversion factors. Terminology was adequate for the targeted group of users (patients or medical professionals) in the majority of cases. The security of personal information used by the apps is assessed by the definition of the privacy policy statement by the developer. As seen in Table 2, more than half of apps did not have a declared privacy policy. However, if we consider the necessity of web access to function as one of the indicators of security of information, only 10% of apps were in need of it. However, when downloading, 14 downloaded apps (27%) required access to one or several of the following mobile phone information on: location, wi-fi connection, phone, photos/media/files, camera, microphone, device ID, and call.

Statistical analysis

Spearman's correlation analysis of relationships among user's ratings found in the app stores, quality elements defined by the MARS score, technical aspects of the app, as well as the affiliation of the developer are presented in Table 3. All MARS individual scoring items were significantly and positively associated with each other, as well as with the Subjective quality and Number of technical aspects (except the Functionality and Number of technical aspects, where there was no significant correlation). Also, the total MARS score and individual items were significantly correlated with the affiliation of the app developers, as expected. However, there were no significant correlations between the ratings from the app stores and the MARS scores. Users' ratings were only in significant

Table 1: Apps included in the evaluation process, n = 52.

App	Developer	Software platform
1) Apps that offer medical advice about symptoms and health queries with the possibility to upload laboratory test results, which can be seen, stored and shared (5/52, 9.6%)		
Babylon Health Online Doctor	Babylon Health, London, UK	Android, iPhone
Biolab	Biolab, Amman, Jordan	Android, iPhone
Blood Glucose Tracker	Little Bytes Software, Albuquerque, USA	Android
Medical Records Assistant	Shadi AboelNaga	Android
My Health	Vaakhayam Technologies Private Limited, Bangalore, India	Android
2) Reference ranges of selected analysis with basic information about the causes of increase or decrease designed for patients (19/52, 36.5%)		
Blood Test Guide	O'Clock Software Pvt, Ltd	Android, iPhone
Blood Test Guide	Rola Tech	Android
Blood Test Result	Knowledge4world	Android
Blood Tests Results Explained	Vilov Frisky Apps	Android
Blood Test Guidance for Perfect Diagnosis	GIF Developer	Android
Blood Tests Results Explained	Brain 2016	Android
Blood Tests Results	Dondur	Android
Blood Test Result	Alpesh Patel	Android, iPhone
Blood Test Result	King of Story App	Android
CALIPER App	The Hospital for Sick Children, Toronto, Canada	Android
Lab Values Free	Sammysoft	Android
Lab Values With Interpretation	HS Developers	Android
Laboratories	Zedney Creative, Jeddah, Saudi Arabia	Android, iPhone
Normal Lab Values	Nooglesoft, New Delhi, India	Android
Pathology Lab Tests Guide – Blood Test Result Guide	Informatic App	Android
Pathology Lab Dictionary	Akshar Clearing Agency	Android
Quick LabRef	Nika Informatics	Android
Mini LAB	Ronny Makhoul	iPhone
Smart Medical Reference	Imago LLC	Android, iPhone
3) Quick reference for laboratory tests for medical students and doctors (20/52, 38.5%)		
Docty – Medical & Laboratory Reference Values	Diego Capela	Android
Lab Gear – Medical Lab Tests	Medical Gear	Android, iPhone
Lab Test Reference Range (free & offline)	Rishikeshav Acharya	Android
Lab Values Reference (Free)	Y Usha Reddy	Android
Lab Values Reference	Ajay Sharma Modgil	Android
MedLab Tutor	Augusta University, USA	Android, iPhone
Normal Lab Values Reference	Nooglesoft, New Delhi, India	Android
Lab Results	Polwarth Medical Ltd	iPhone
Medical Lab Tests	Medicon Apps	iPhone
Davis's Lab & Diagnostic Tests	Unbound Medicine, Inc	Android, iPhone
Skyscape Lab Values Medical Reference	Skyscape Medpresso Inc	Android, iPhone
Lab Essentials – Medical Lab Values	Med Gears	Android, iPhone
Pagana: Diagnostic and Lab Tests	Unbound Medicine, Inc	Android, iPhone
Lab Values Reference Guide	William Wyatt	iPhone

Table 1 (continued)

App	Developer	Software platform
Lab Test Navigator 800+	ClinLab Navigator, LLC	iPhone
qLab	Intern 911 Medical Applications, LLC	iPhone
Lab Notes: Guide to Lab & Diagnostic Tests	Skyscape Medpresso Inc	Android, iPhone
MA Notes: Medical Assistant's Pocket Guide	Skyscape Medpresso Inc	Android, iPhone
Fischbach's Nurse's Manual of Laboratory and Diagnostic Tests	Skyscape Medpresso Inc	Android, iPhone
Bakerman's ABC's of Interpretive Laboratory Data	Skyscape Medpresso Inc	Android, iPhone
4) Apps for monitoring the state of user's health through a wide range of health parameters, including glucose and/or cholesterol as laboratory data (7/52, 13.5%)		
Health Assistant	Mromso & Friends	Android
Health Diary	PNN Soft, Kyiv, Ukraine	Android
My Health Tracker	Walter Gross	Android
S Health	Samsung Electronics Co, South Korea	Android
Tactio Health	Tactio Health Group	Android
Fisike	Healthcare assistant	iPhone
HealthKey	Key Management Group	iPhone
5) Apps that provide access to patients' laboratory results to physicians (0/52, 0%)		
6) Apps that enable patients to access their laboratory test results directly from the diagnostic center (1/52, 1.9%)		
Lab Reports	Hellenic Informatics LTD	iPhone
7) Electronic health records apps that include laboratory test results (0/52, 0%)		

correlation with the corresponding number of ratings in each app store.

According to Kruskal-Wallis analysis (Table 4), there was a significant difference in the number of technical aspects employed, MARS Engagement and Informational quality score items, MARS Total score, and MARS subjective quality. The lowest values for all of these items were in the category 2, i.e. the category of apps that provided reference ranges of selected analysis with basic information about the causes of increase or decrease designed for patients. Moreover, the post hoc test showed that the difference was statistically significant between the values in category 2, and the values in all other categories of apps. Affiliation of app providers was in no connection with these differences, as the χ^2 independence test showed independent distribution of affiliation categories and app classification with $p = 0.3414$ (data not shown).

Discussion

To the best of our knowledge, the analysis of the number and quality of available smartphone apps using laboratory medicine data in any way, has not yet been performed. These mobile apps are not subjected to regulatory oversight as they are not considered medical devices [16]. Considering the widespread use of smartphones and the availability of mobile apps, increasing impact of their use and the information they provide on overall health related behavior is expected. Therefore, we wanted to evaluate the quality and reliability of information regarding laboratory medicine data in mobile apps present on the app market.

As seen from the presented data, laboratory medicine data present in any way in mobile apps on the Android and iOS market are represented in a minor percent. This number is much lower than, for example, for apps for the prevention, detection, and management of cancer, where out of the total number of 1314 applications, 295 met the selection criteria [17]. However, the number of apps that are dealing exclusively with laboratory medicine data and their interpretation is half of the total number (categories 2 and 3), and only half of these are intended for patient use. When searching PubMed, we found no articles that evaluated any of applications dealing with laboratory medicine data, which is why we have no data on their actual use and effect on patients' habits and overall health.

Although several possible criteria can be found for mobile apps content evaluation throughout the literature, we used a multidimensional tool for classifying and rating the quality of mobile health apps, the MARS app quality

Table 2: Descriptive data for the analyzed apps (n=52).

	n (%)	Mean (SD)	Median (IQR)	Range
App basic descriptive information				
Average user rating on GP (scale 1–5)	33 (78%)	4.2 (0.4)	4.2 (0.5)	3.3–5.0
Average user rating on iTunes (scale 1–5)	11 (44%)	4.3 (1.0)	4.7 (1.0)	1.7–5.0
Number of rating on GP (count)	33 (78%)	702 (1939)	61 (176)	0–8852
Number of rating on iTunes (count)	11 (44%)	29 (43)	8 (22)	2–138
Price on GP (€)	10 (28%)	38.43 (23.16)	45.99 (29.80)	0.99–64.99
Price on iTunes (\$)	12 (48%)	36.90 (23.49)	45.99 (37.42)	1.99–64.99
Last update (year)	–	2016 (1.3)	2017 (2)	2013–2018
Affiliations				
Unknown	18 (34.6%)	–	–	–
Commercial	30 (57.7%)	–	–	–
Government	0 (0%)	–	–	–
NGO	1 (1.9%)	–	–	–
University	3 (5.8%)	–	–	–
Category in the store				
Medical	37 (71.2%)	–	–	–
Health and fitness	11 (21.2%)	–	–	–
Education	3 (5.8%)	–	–	–
Lifestyle	1 (0.5%)	–	–	–
Technical aspects				
Number of technical aspects per app (0–6)	52 (100%)	0.98 (1.39)	0.00 (2.0)	0–5
Presence of technical aspects				
Allows sharing	17 (33%)	–	–	–
Has an app community	4 (8%)	–	–	–
Requires login	10 (19%)	–	–	–
Allows password protection	5 (10%)	–	–	–
Sends reminders	8 (15%)	–	–	–
Needs web access to function	5 (10%)	–	–	–
MARS app quality rating (1–5)				
Engagement	52 (100%)	2.67 (0.85)	2.6 (1.2)	1.2–4.6
Functionality	52 (100%)	4.89 (0.28)	5.0 (0.0)	3.5–5.0
Aesthetics	52 (100%)	3.78 (0.60)	3.7 (0.7)	2.7–5.0
Information quality	52 (100%)	3.57 (0.83)	3.8 (1.0)	1.6–4.8
Total score	52 (100%)	3.73 (0.55)	3.8 (0.8)	2.2–4.8
Subjective quality	52 (100%)	3.21 (0.79)	3.2 (1.0)	1.2–4.8
Laboratory medicine specific features				
Comparability with locally accepted reference ranges				
Yes	38 (73%)	–	–	–
No	9 (17%)	–	–	–
Unit of concentration				
SI	1 (1.9%)	–	–	–
Conventional	24 (46.2%)	–	–	–
Both	23 (44.2%)	–	–	–
Appropriate terminology used				
Yes	47 (90.4%)	–	–	–
No	2 (3.8%)	–	–	–
Declared privacy policy for user's data				
Yes	21 (40.4%)	–	–	–
No	28 (53.8%)	–	–	–
Interpretative elements consistent with best practice				
Yes	29 (55.8%)	–	–	–
No	7 (13.5%)	–	–	–

SD, standard deviation; IQR, interquartile range; GP, Google Play; NGO, non-governmental organization; MARS, Mobile Application Rating Scale.

Table 3: Spearman's ρ between number of technical aspects present, MARS scores, users' ratings, and producers' affiliations of the reviewed apps.

	1	2	3	4	5	6	7	8	9	10	11	12
1. Number of technical aspects	1.00											
2. MARS engagement	0.542 ^a	1.00										
3. MARS functionality	0.094	0.345 ^b	1.00									
4. MARS aesthetics	0.283 ^b	0.594 ^a	0.286 ^b	1.00								
5. MARS information quality	0.334 ^b	0.701 ^a	0.468 ^a	0.702 ^a	1.00							
6. MARS total score	0.433 ^a	0.874 ^a	0.478 ^a	0.819 ^a	0.911 ^a	1.00						
7. MARS subjective quality	0.348 ^b	0.634 ^a	0.485 ^a	0.579 ^a	0.846 ^a	0.802 ^a	1.00					
8. Average rating GP	-0.101	-0.173	-0.173	-0.057	-0.035	-0.110	0.085	1.00				
9. Average rating iTunes	0.092	-0.061	0.157	0.355	0.186	0.160	0.320	0.064	1.00			
10. Number of ratings GP	0.139	0.077	0.109	0.084	0.062	0.082	0.234	0.517 ^a	0.152 ^a	1.00		
11. Number of ratings iTunes	0.120	-0.029	0.048	0.314	0.076	0.075	0.245	0.236	0.891 ^a	0.283	1.00	
12. Affiliations	0.041	0.429 ^a	0.366 ^a	0.347 ^a	0.496 ^a	0.514 ^a	0.416 ^a	-0.097	0.032	0.009	0.068	1.00

GP, Google Play; MARS, Mobile Application Rating Scale. ^a $p < 0.01$; ^b $p < 0.05$.

criteria. The MARS rating was validated on weight management and mental health apps [12, 18], and was recommended by Grundy et al. [10] as the innovative method in the study of commercially available health-related apps. The average MARS total score was above average, with the value of 3.8 out of 5. The highest individual MARS scoring item was for the Functionality 5.0, and the lowest was Engagement, with median individual score value of 2.6. The analysis of MARS score values between the seven different categories of applications that use laboratory medicine data, revealed the poorest performance and quality of the apps intended for patients. This, together with the significant issues of security of personal information used by the apps, and the questionable affiliation of developers, which was either commercial or unknown, without referencing the source of information cited in the app, reveals the poor quality of apps aimed to be the source of information to patients who want to know more about their laboratory test results. Furthermore, there were no significant correlations between the ratings from the app stores and the MARS scores. The users ratings from the app store could be misleading as they are unlikely to have enough knowledge in health care, and also often among reviews contributing ratings there are those made by services hired by publishers to provide positive ratings and reviews for their apps [19].

Limitations of this analysis are that we have not evaluated the impact these apps have on their intended users, i.e. patients and medical professionals. It is important to identify whether the use of these apps leads to a change in behavior – increased health awareness, understanding of laboratory test results, and their impact on personal health. Also, we used the MARS rating for apps quality assessment, a general tool not yet evaluated on laboratory medicine apps. As there is a chance that app developers restrict distribution of apps to specific countries, their availability for download may depend on the location of the user. Therefore, the fact that the app search was conducted within Google Play and iTunes in Serbia may be considered as a limitation of this study. Further work in this direction would provide the definition of quality benchmarks for laboratory medicine apps, and possible modification of the MARS items for their adequate evaluation.

In conclusion, mobile apps for laboratory medicine available on the app market that deal exclusively with the interpretation of laboratory test results represent half of the apps selected with the defined search criteria. Apps designed for patients, the vulnerable group, are of the poorest quality, considering the total quality of the content and information they provide, estimated using

Table 4: Kruskal-Wallis test of differences between number of technical aspects present, MARS score items, and users' ratings of apps in the seven distinguished classes of apps using laboratory medicine data.

App classes	Median (IQR)							p-Value ^a
	1	2	3	4	5	6	7	
	n=5	n=19	n=20	n=7	n=0	n=1	n=0	
Number of technical aspects	4.0 (1.5)	0.0 (1.0) ^b	0 (0.5) ^b	2.0 (1.25) ^{c,d}	/	5.0 (0.0) ^{c,d}	/	<0.0001
MARS								
Engagement	3.15 (0.8)	2.0 (0.9) ^b	2.6 (0.6) ^{b,c}	3.6 (1.0) ^{c,d}	/	3.6 (0.0) ^c	/	<0.0001
Functionality	5.0 (0.0)	5.0 (0.25)	5.0 (0.0)	5.0 (0.06)	/	5.0 (0.0)	/	0.0782
Aesthetics	3.7 (1.4)	3.3 (0.6)	3.7 (0.7)	4.0 (1.2)	/	5.0 (0.0)	/	0.1753
Informational quality	4.2 (0.8)	2.8 (1.5) ^b	3.75 (0.55) ^c	4.2 (0.3) ^c	/	4.5 (0.0) ^c	/	0.0034
Total score	4.2 (0.7)	3.2 (0.8) ^b	3.8 (0.4) ^{b,c}	4.25 (0.6) ^{c,d}	/	4.5 (0.0) ^c	/	0.0009
MARS subjective quality	4.0 (1.25)	2.75 (1.4) ^b	3.25 (0.75)	3.5 (0.3) ^c	/	4.25 (0.0) ^c	/	0.0129
Average rating GP	4.1 (0.4)	4.25 (0.5)	4.3 (0.6)	3.6 (0.5)	/	/	/	0.2162
Average rating iTunes	4.4 (0.9)	4.35 (1.3)	4.6 (1.0)	/	/	5.0 (0.0)	/	0.6392
Number of ratings GP	184 (3729)	32 (177)	60 (211)	135 (130)	/	/	/	0.7185
Number of ratings iTunes	73 (130)	16 (15)	7 (23)	/	/	2 (0)	/	0.2851

1, Apps that offer medical advice about symptoms and health queries with the possibility to upload laboratory test results, which can be seen, stored and shared; 2, reference ranges of selected analysis with basic information about the causes of increase or decrease designed for patients; 3, quick reference for laboratory tests for medical students and doctors; 4, apps for monitoring the state of user's health through a wide range of health parameters, including glucose and/or cholesterol as laboratory data; 5, physicians' access to patients' results directly from the providing laboratory; 6, apps that enable patients to access their laboratory test results directly from the diagnostic center; 7, Mobile Electronic Health Record (EHR) apps; IQR, interquartile range; MARS, mobile application rating scale; GP, Google Play; ^aKruskal-Wallis test, statistically significant difference $p < 0.05$ (bolded); ^bpost hoc analysis, statistically significant difference ($p < 0.05$) with the value in the class 1; ^cpost hoc analysis, significant difference ($p < 0.05$) with the value in the class 2; ^dpost hoc analysis, significant difference ($p < 0.05$) with the value in the class 3.

the MARS tool. This estimation needs to be validated for laboratory medicine apps, and eventually modified after consideration of specific quality benchmarks.

Author contributions: All the authors have accepted responsibility for the entire content of this submitted manuscript and approved submission.

Research funding: None declared.

Employment or leadership: None declared.

Honorarium: None declared.

Competing interests: The funding organization(s) played no role in the study design; in the collection, analysis, and interpretation of data; in the writing of the report; or in the decision to submit the report for publication.

References

- McNutt RA. Shared medical decision making problems, process, and progress. *J Am Med Assoc* 2004;292:2516–8.
- Hallworth MJ, Epner PL, Ebert C, Fantz CR, Faye SA, Higgins TN, et al. on behalf of the IFCC Task Force on the Impact of Laboratory Medicine on Clinical Management and Outcomes. Current evidence and future perspectives on the effective practice of patient-centered laboratory medicine. *Clin Chem* 2015;61: 589–99.
- Watson ID, Oosterhuis WP, Jorgensen PE, Dikmen ZG, Siodmiak J, Jovicic S, et al. A survey of patients' views from eight European countries of interpretive support from specialists in laboratory Medicine. *Clin Chem Lab Med* 2017;55:1496–500.
- Aungst TD, Clauson KA, Misra S, Lewis TL, Husain I. How to identify, assess, and utilize mobile medical applications in clinical practice. *Int J Clin Pract* 2014;68:155–62.
- Risk A, Dzenowagis J. Review of internet health information quality initiatives. *J Med Internet Res* 2001;3:e28.
- Gagliardi A, Jadad AR. Examination of instruments used to rate quality of health information on the internet: Chronicle of a voyage with an unclear destination. *Br Med J* 2002;324:569–73.
- US National Library of Medicine Evaluating Internet Health Information: A tutorial from the National Library of Medicine 2016. <http://medlineplus.gov/webeval/webeval.html>. Accessed 3 Jun 2018.
- Kharazzi H, Chisholm R, VanNasdale D, Thomson B. Mobile personal health records: an evaluation of features and functionality. *Int J Inform* 2012;81:579–93.
- Xie B, Su Z, Zhang W, Cai R. Chinese cardiovascular disease mobile apps' information types, information quality, and interactive functions for self-management: systematic review. *JMIR Mhealth Uhealth* 2017;5:e195.
- 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;51:1051–9.

11. BinDhim NF, Hawkey A, Trevena L. A systematic review of quality assessment methods for smartphone health apps. *Telemed e-Health* 2015;21:97–104.
 12. Stoyanov SR, Hides L, Kavanagh DJ, Zelenko O, Tjodronegoro D, Mani M. Mobile App Rating Scale: a new tool for assessing the quality of health mobile apps. *JMIR mHealth Uhealth* 2015;3:e27.
 13. O'Neill S, Brady RR. Colorectal smart phone apps: opportunities and risks. *Colorectal Disease* 2012;14:e530–4.
 14. Hallgren KA. Computing inter-rater reliability for observational data: an overview and tutorial. *Tutor Quant Methods Psychol* 2012;8:23–34.
 15. Koo TK, Li MY. A guideline of selecting and reporting intraclass correlation coefficients for reliability research. *J Chiropr Med* 2016;15:155–63.
 16. US Food and Drug Administration (FDA). Mobile medical applications: Guidance for industry and food and drug administration staff. Silver Spring, MD: US FDA:2015.
 17. Bender JL, Yue RY, To MJ, Deacken L, Jadad AR. A lot of action, but not in the right direction: systematic review and content analysis of smart phone applications for the prevention, detection, and management of cancer. *J Med Internet Res* 2013;15:e287.
 18. Bardus M, van Beurden SB, Smith JR, Abraham C. A review and content analysis of engagement, functionality, aesthetics, information quality, and change techniques in the most popular commercial apps for weight management. *Int J Behav Nutr Phys Act* 2016;13:35.
 19. App review services. Available at www.appreviewservice.com/. Accessed 2 Jun 2018.
-
- Article note:** Lecture given by Dr. Snežana Jovičić at the 2nd EFLM Strategic Conference, 18–19 June 2018 in Mannheim (Germany) (<https://elearning.eflm.eu/course/view.php?id=38>).