

RESEARCH ARTICLE

ACCU³RATE: A mobile health application rating scale based on user reviewsMilon Biswas¹, Marzia Hoque Tania², M. Shamim Kaiser^{3*}, Russell Kabir⁴, Mufti Mahmud⁵, Atika Ahmad Kemal⁶

1 Computer Science and Engineering, Bangladesh University of Business and Technology, Mirpur, Dhaka, Bangladesh, **2** Department of Engineering Science, Institute of Biomedical Engineering, University of Oxford, Oxford, United Kingdom, **3** Institute of Information Technology, Jahangirnagar University, Savar, Dhaka, Bangladesh, **4** School of Allied Health, Faculty of Health, Education, Medicine and Social Care, Chelmsford, United Kingdom, **5** Department of Computer Science, Nottingham Trent University, Nottingham, United Kingdom, **6** Management and Marketing at Essex Business School (EBS), University of Essex, Colchester, United Kingdom

☞ These authors contributed equally to this work.

* mskaiser@juniv.edu



OPEN ACCESS

Citation: Biswas M, Tania MH, Kaiser MS, Kabir R, Mahmud M, Kemal AA (2021) ACCU³RATE: A mobile health application rating scale based on user reviews. PLoS ONE 16(12): e0258050. <https://doi.org/10.1371/journal.pone.0258050>

Editor: Junaid Rashid, Kongju National University, South Korea, KOREA, REPUBLIC OF

Received: April 16, 2021

Accepted: September 13, 2021

Published: December 16, 2021

Peer Review History: PLOS recognizes the benefits of transparency in the peer review process; therefore, we enable the publication of all of the content of peer review and author responses alongside final, published articles. The editorial history of this article is available here: <https://doi.org/10.1371/journal.pone.0258050>

Copyright: © 2021 Biswas et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability Statement: All data files used in this study are available from the Dryad database (data link: <https://doi.org/10.5061/dryad.jdfn2z3bf>).

Funding: The author(s) received no specific funding for this work.

Abstract

Background

Over the last decade, mobile health applications (mHealth App) have evolved exponentially to assess and support our health and well-being.

Objective

This paper presents an Artificial Intelligence (AI)-enabled mHealth app rating tool, called ACCU³RATE, which takes multidimensional measures such as user star rating, user review and features declared by the developer to generate the rating of an app. However, currently, there is very little conceptual understanding on how user reviews affect app rating from a multi-dimensional perspective. This study applies AI-based text mining technique to develop more comprehensive understanding of user feedback based on several important factors, determining the mHealth app ratings.

Method

Based on the literature, six variables were identified that influence the mHealth app rating scale. These factors are user star rating, user text review, user interface (UI) design, functionality, security and privacy, and clinical approval. Natural Language Toolkit package is used for interpreting text and to identify the App users' sentiment. Additional considerations were accessibility, protection and privacy, UI design for people living with physical disability. Moreover, the details of clinical approval, if exists, were taken from the developer's statement. Finally, we fused all the inputs using fuzzy logic to calculate the new app rating score.

Results and conclusions

ACCU³RATE concentrates on heart related Apps found in the play store and App gallery. The findings indicate the efficacy of the proposed method as opposed to the current device

Competing interests: The authors have declared that no competing interests exist.

scale. This study has implications for both App developers and consumers who are using mHealth Apps to monitor and track their health. The performance evaluation shows that the proposed mHealth scale has shown excellent reliability as well as internal consistency of the scale, and high inter-rater reliability index. It has also been noticed that the fuzzy based rating scale, as in ACCU³RATE, matches more closely to the rating performed by experts.

Introduction

The world population is rising rapidly which is causing the present healthcare facilities to fall short in meeting the need. To cope to this need, many countries are reforming their healthcare provisioning related to assistive technologies which is a major move aiming to integrate clinical care and consumer health [1, 2]. However, there is a clear imbalance in the supply and demand of healthcare facilities in majority of the countries. This imbalance is not only seen in developing countries, but also in developed ones during high demand scenarios such as the current novel coronavirus disease (COVID-19) pandemic [3, 4]. In addition to insufficient availability of lower-level government healthcare facilities, in a multi-tier healthcare delivery system it's appropriate functioning is severely hampered by inaccessibility to secondary and tertiary healthcare services [5].

In recent years, the massive deployment of the information and communication technology (ICT) infrastructure and enormous improvement of tele-density along with availability of relatively inexpensive smart gadgets, mobile-based therapies (called mHealth) have aimed to confront these issues [6, 7]. These mHealth apps have facilitated medical research and practice in public health through the usage of mobile devices such as tablets and smartphones [8–10]. Paralelly, since the last decade, “patient-centric healthcare” has been gaining popularity in delivering personalised healthcare [11–14]. The technological advances in healthcare provisioning, such as development of mHealth and various portable Health (pHealth) app, allow people to make informed decisions about medical services to monitor their condition, and recover faster from illnesses [15–17]. An mHealth-based patient-centric healthcare approach includes personalised care facilities to meet the demands of the individual needs of a patient. The mHealth app makes it easier for patients to gain quicker access to medical care, enables real-time and continuous patient monitoring, and facilitates medication intake accuracy to increase patient safety [18]. This has been applied to promotes self-management of many chronic illnesses such as diabetes, hypertension and cardiovascular disease [19–22]. This increasing trend in the usage of mHealth apps by consumers has attracted a significant portion of academic research in understanding consumer experience, decision making and behaviour in selecting a pertinent mHealth app over similar alternatives [23–25]. As a result, it's important to understand that mHealth apps are based on evidence and behavioural research in order to encourage long-term use and have a significant impact on health outcomes [10, 26, 27].

Initially, research in the “digital healthcare industry” focused on defining and cataloguing mobile app functionality and characteristics [28, 29]. However, if consumers are expected to use mHealth apps on a consistent basis in order to engage in self-management of their own well-being, it is critical to obtain a deeper understanding of the characteristics and methods that contribute to long-term app usage decision making [10, 30–32]. It is easily noticeable that quite some efforts have been put to address the need of consumers in a healthcare app. Approximately over 100,000 apps are available in the Apple iTunes Store and Google Play Store pertaining to categories such as medical, health and fitness [8]. Recent research shows

that an average smartphone user may have around 80 apps installed on their phone, however, usage of these apps is restricted to only 9 apps per day [33].

The motivation for this paper arises from the very little conceptual understanding that is present on how users' online reviews determine an app's rating, particularly, in the context of mHealth apps. Studies on mHealth apps primarily focus on users' long term set of experiences [10], theorise how cognitive factors impact users' continued intention of using mHealth apps [34] and examine factors influencing the download of these apps [35]. Previous research shows that an app's features and characteristics make positive contributions to the app's rating, thus it is worthwhile to investigate further the rating mechanism of mHealth apps [1]. A set of multidimensional features may better explain a large portion of an app's rating. Hence, it is imperative to gain richer insights into the apps available in stores and quantify the quality of each app's feature through evaluation of user reviews using artificial intelligence (AI)-enabled text mining methods.

Moreover, the limited AI-led research operationalizes users' reviews for assigning the rating of mHealth app. Various dimensions may represent the propensity of a review in attracting users to download a specific app, as user reviews are generally known to be informative and help users finding the most appropriate one. Thus, for a variety of stakeholders (e.g., healthcare providers, app developers and patients) in the healthcare industry, it is pertinent to gain deeper insights on any app's functionality and characteristics from multiple perspectives that have been extracted through text mining of user reviews. Furthermore, AI is an emerging field of research that focuses more on detailed text information generated in reviews. Previous research has used content analysis which is very time consuming, especially when the amount of textual data is very high [1, 35–37]. Text mining has considerable amount of popularity leading to different techniques to quantify textual content [38–41]. While research has evidenced mixed results on the influence of semantic characteristics, other scholars have examined the effects of textual characteristics in online user reviews for helpfulness and effectiveness in online consumer retail markets [40].

However, there is no consistent evidence of how text mining user reviews may distinctly influence mHealth app rating based on multidimensional measures. Thus, it becomes imperative to evaluate the effects of textual characteristics in determining user reviews for mHealth apps. We draw on this knowledge gap that motivates this study and sets it apart from previous studies. We explore the effects of various characteristics of online user reviews and carry out sentiment analysis of mHealth apps which may then assign a new rating, substantially different from the original rating. In the process of creating this new rating, in addition to the existing user inputs (*user star rating*, *user text review*), there are other important factors, such as, *user interface (UI) design*, *functionality*, *security and privacy*, and *clinical approval and certification*, also needs to be taken into consideration. In view of all these factors, this paper presents an AI-enabled mHealth app rating tool, named, ACCU³RATE (clinical Approval & CertifiCation, fUnctionality, Ui degisn, User Rating privAcY & securiTy, and user rEview), based on six factors that influence the mHealth app rating scale. However, considerations such as functionality, security and privacy, UI design and clinical approval details were taken from the developer statement and details provided by the app store.

Within context of the current work, fuzzy logic has been used in determining the relationship among the features. It is highly elastic and is able to incorporate the views of experts into the model in the form of IF-THEN statements. Jumbled, imprecise, inaccurate, and erroneous knowledge can be inputted to the fuzzy system. Thus, a fuzzy logic knowledge fusion technique can be utilized to integrate the effect of all factors mentioned above and determines the new app rating value for the target mHealth app. Our findings suggest that the semantic feature of reviews have a significant impact on the rating score of the mHealth app.

The major contributions from this paper can be summarized as–

- We describe AI-enabled mHealth app rating tool, called ACCU³RATE, which takes multidimensional measures such as user star rating, user review and features declared by the developer to generate the rating of an app.
- We present AI based text mining and sentiment analysis to evaluate the users review for selected apps.
- We introduce a fuzzy logic based knowledge fusion technique that determines the score of the mHealth app based on input factors.
- The results show the effectiveness of selected factors in the rating score of the mHealth app.

The paper is structured as follows. Section provides a theoretical review of mHealth applications and determinants that affect mobile rating. Section describes basics of mHealth app and factors affecting the mHealth app rating scale. In Section, we propose the AI-enabled app rating model. Section shows results and discussions, and finally the work is concluded in Section, followed future direction of this research.

Literature review

Overview of mobile health applications

Over the last few years, we have observed an excessive evolution of mobile health applications than any other innovation in health care [42]. The term ‘app’ is an abbreviation of the word ‘application’ that refers to a program that has been developed for a specific purpose and is generally configured to run on handheld devices like cell phones, tablets, computers, and certain wearable devices such as smartwatches [43]. In recent times, we have observed a proliferation of mobile apps that are mostly created for only two of the most important tech giants of smartphone operating systems- Apple IOS and Google Android. The World Health Organization (WHO) defines these applications as, ‘medical and public health practice supported by mobile devices such as mobile phones, patient monitoring devices, personal digital assistants (PDAs) and other wireless devices’. Similarly, according to Hamel et al. [44] mHealth ‘is the use of portable devices such as smartphones and tablets to improve health’. mHealth apps in practice run on mobile devices with remote network connectivity and dynamic execution of contexts [45]. Context is any information that can be used to characterize the situation of an entity, being either a person, place or object [46]. The context awareness system collects real-time data from patients in a comprehensive manner and presents them to health care professionals to manage their tasks in order to increase the quality of patient care [47, 48].

Extant literature on mHealth applications exhibits how mHealth apps are commonly used for health education, disease self-management, remote monitoring of patients, and collection of dietary data [49, 50]. The information collected through mHealth apps can provide useful support for encouraging health behaviour change, chronic disease self-care and effective management of many conditions while keeping healthcare providers informed of the patient’s condition [50]. This also strengthens the relationship between the patient and the provider [51]. According to Scoping review findings, mHealth apps provide the potential for general practitioners to take medical history and make diagnoses, perform some physical examinations, aid in clinical decision making, and manage long term disease-specific care and promote general health well-being [52]. Whilst traditionally, health care is delivered to individuals through face to face interaction with health care professionals, technology advancement has facilitated communication between patients and health care professionals [53]. Moreover, mobile phone apps are commonly incorporated into the design of health promotion programs both face-to-face

and online [54]. Based on these, mHealth apps are mainly categorised into two categories—lifestyle management (such as fitness, lifestyle modification, diet and nutrition) and chronic disease management apps (such as mental health, diabetes, and cardiovascular diseases). The other categories include- self-diagnosis, meditational re-minders, and patient portal apps [43].

Studies document how mHealth apps provide information and health behaviour interventions at a meagre cost, that is, easy to access and personalized for user-specific needs [55]. A systematic review of mHealth apps for chronic diseases showed that there are improvements in patient lifestyle related to, healthier eating, weight loss, controlled blood pressure and glycaemic levels, including treatment awareness options among patients [21]. With mHealth apps, users can receive important information through push notifications and can monitor their health status at a personal level. Litman and his colleagues [56] from their mobile exercise App research found that exercise app users are more likely to exercise during their leisure time as compared to those who do not use exercise apps. They also highlighted that exercise apps may potentially increase exercise levels and health outcomes such as BMI levels. However, it is also noted that mHealth apps also pose some limitations, and users may face a life-threatening situation if the mHealth apps are not switched to use and made available all the time. In addition, private and sensitive communications are exchanged between the user and the health care provider through these apps. Hence, usability studies should be sufficiently evaluated to determine the practical way in assessing mHealth apps usability [57]. This is critical as any app developer can develop and market a mHealth app through the app store without conducting any usability test [58]. Thus, studies emphasize the importance of a regulatory body to assess the accuracy, quality and performance of health apps, as some mHealth apps were hacked, leading to patient data breaches [43]. Moreover, without any appropriate security measures and processes in place, mHealth app users may become exposed to the severe repercussions of information security and privacy infringements [59]. Hence, we note that there are no general international standards that a data and privacy policy should offer users. This further suggests that, information on how privacy policies determine app star rating is relatively absent. Many privacy policies do not focus on the security features related to the app and therefore do not create confidence in usage for users. In most cases, users may be ignorant and are more likely to use apps which have ambiguous privacy policies [60].

Fig 1(A) shows number of mHealth apps uploaded in the app store each year; Smart phone penetration in the last five years is illustrated in Fig 1(B). The pie chart shows various categories of apps in terms of percentage (See Fig 1(C))

We have found five scales, such as THESIS [42], MARS [61, 62], Brief DISCERN [63], uMARS [64] and ORCHA-24 [65], THESIS [42] is an open-source application rating tool that

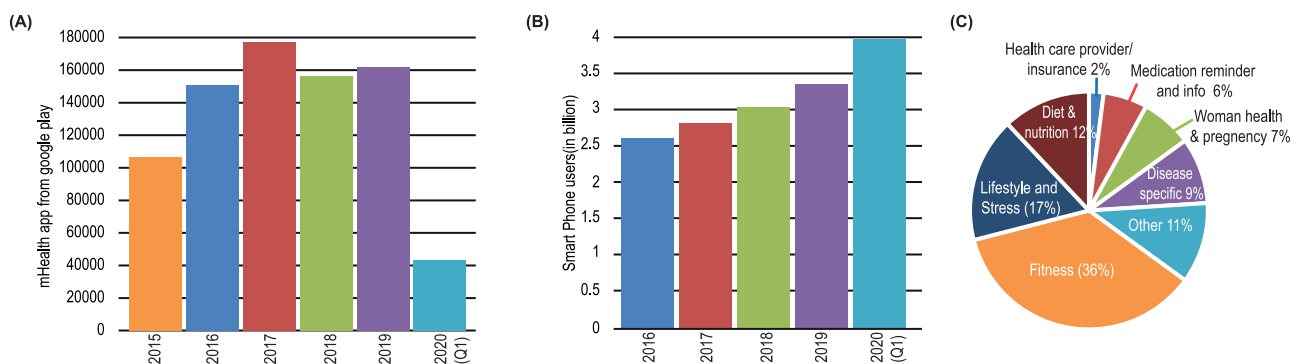


Fig 1. (A) Year-wise number of mHealth apps uploaded in the google play, and (B) Smart phone penetration in the last five years, and (C) Pie chart shows various categories of apps in percentage.

<https://doi.org/10.1371/journal.pone.0258050.g001>

is performed using a Delphi process. The tool can calculate usability, risks, and health app benefits. Authors used a standardized method for selecting chronic disease applications with 4 stars and < 4 stars, then scored them with THESIS for their reliability and internal consistency of the tool. The THESIS testing indicates that applications serving chronic illness patients can perform better, particularly in terms of privacy/safety and interoperability. THESIS guarantees further testing and can lead software and policymakers to continue improving app efficiency, enabling applications to enhance patient results more reliably.

Mobile app rating scale (MARS) [66] is an app rating scale for mHealth that has been developed to measure the reliability and efficiency of mHealth applications. The MARS Scale uses four objective quality scales: “functionality”, “aesthetics”, “engagement” and “information quality”; and one subjective quality scale; which were refined into the “23-item MARS”. The MARS has shown excellent intra-class correlation accuracy and reliability.

The Brief DISCERN score [63] can assess quality of mHealth apps for panic disorder (PD) using factors such as transparency, interactivity, self-help score, and evidence-based content. The Brief DISCERN score showed a number of linear regressions that considerably predicted the content quality and self-help performance.

An end-user version of the MARS, called uMARS [64], proposed to assess the quality of mHealth apps. The uMARS had outstanding internal consistency for all subscales with strong individual alphas. The uMARS are used to accumulate app reviews at 1-, 3-, and 6-month intervals. The uMARS had strong individual alphas for all subscales and had outstanding internal stability.

The Organization for the Review of Care and Health Applications- 24 Question Assessment (ORCHA-24) [65] is National Health Service (NHS) approved scale which can evaluate the quality of applications for Chronic Insomnia Disorder on the Android Google Play store for high quality and low-risk health apps. The ORCHA-21 system performed evaluation using evaluation criterion such as efficiency; engagement; protection and privacy; user interface, and quality of clinical efficacy. Table 1 shows the various mHealth app rating scales found in the literature.

None of the rating scales considered, to the best of our knowledge, various characteristics of online user feedback and performed sentiment analysis for mHealth applications, which could then be assigned a new rating, substantially different from the original rating. Thus, the proposed mHealth app rating tool, called ACCU³RATE, uses evaluation criteria such as user star rating, user text review, UI design, functionality, security and privacy, and clinical approval.

Table 1. Various mHealth app rating scale.

Ref.	Scale Name	Factors	Performance Metric	Results
[42]	THESIS	U; S/P; THC; T	IRR	$k > 0.31$
			ICC	$\rho_T = 0.85$
[66]	MARS	E; F; Es; I; Q	ICC	$\rho_T = 0.79$
[63]	Brief DISCERN	T; Q	LR	$R^2 \approx 0.7$
[64]	uMARS	E; F; Es; I; Q	ICC	$\rho_T = 0.90$
[65]	ORCHA-24	Dg; Cea; UX	IRR	$k > 0.88$
			ICC	$0.84 \leq \rho_T \leq 0.93$

In Factors Column: U–Usability; S/P–Security/privacy; THC–Technical and Health Content; T–Transparency; E–Engagement; F–Functionality; Es–Esthetics; I–Information; Q–Quality; Dg–Data governance; Cea–Clinical efficacy and assurance; UX–User experience and engagement;

In Performance Metric Column: IRR–Interrater reliability; ICC–Intra-class Co-relation Co-efficient; LR–Linear Regression;

<https://doi.org/10.1371/journal.pone.0258050.t001>

Determinants of mobile health app ratings

At present, there is little known about which mHealth apps are most commonly used and whether they align with app quality [67]. Majority of the mHealth apps found in the app stores are available in trial versions to attract potential users or as a promotional strategy to increase the adoption rate [68]. Fiordelli et al. [9] in their systematic review brought to our attention that there are millions of mHealth apps (free and paid) publicly available on app stores, and to-date they have not been evaluated them. Byambasuren et al. [58] highlighted that low quality and lack of effectiveness hinders the usability of mHealth apps, as the majority are developed by non-health care organisations which raises questions about the accuracy and trustworthiness of these apps [69].

While evaluating behavioral mHealth apps, studies report that app ratings are broadly inconsistent and often contradictory while most popular behavioural health apps rating are not favorable [67]. Furthermore, other frequently downloaded behavioral mHealth apps have questionable support in the literature and offer no evidence-based behavioural support. For example, research on the usage of mHealth apps by cancer survivor patients revealed that some apps used theoretical models of behaviour change, but the majority of the apps in the market did not apply any of those theoretical elements in their apps [55].

Some scholars suggest that independent and reliable sources should evaluate mHealth apps and should recommend a collection of trustworthy apps to health care professionals to, be referred to patients [58]. According to Levine et al. [42] an app rating tool can be developed to identify mHealth apps that may cause potential harm and breach the privacy policy. When an appropriate mHealth app is prescribed by health care professionals, they must be confident about the functionality of the app, maintenance of data privacy and its usability [58]. A survey conducted among mHealth app users in the USA found that 82 percent users would change providers if they were aware of alternative apps which were more secure [70]. There is also no clear evidence available to suggest that an app developed by a health care organization will have more usability. Bivji et al. [69] identified that price, user ratings, in-app purchase options and in-app advertisements were the greatest predictors of app downloads from the app store.

Other studies present that the actual usability and quality rating scales are primarily developed only for professionals, and not for end-users who are patients or care providers. It is widely recognized that rating scales that are usable by all end users would make mHealth apps more accessible and meaningful to consumers [71]. Janatkah et al. [72] found that a mHealth app's usability is associated with the education level, employment and place of user's residence. While there is evidence of a weak relationship between user app ratings and usability and clinical utility of the app [73] in the literature, generally we note that users make decisions related to app use, based on the title, price, star ratings, reviews and number of downloads [74]. Furthermore, mHealth users disproportionally favour apps that take advantage of unique features of smartphones with higher ratings and present the real benefits to users to keep track of their health records [45].

Generally, we observe that mHealth app users' opinions or satisfaction levels, star rating system, app description of checklists are some common features used to measure the quality of the apps. However, scholars argue that none of these methods are scientifically adequate to measure the quality of the apps [75]. Singh et al. [73] identified that store rating alone is not enough to determine whether an app is of high quality but inclusion of clinicians and users' reviews will ensure that the recommended apps are usable and clinically useful.

Further, we note that in the literature users' reviews of mHealth apps in relation to app rating is generally scarce. There is research on the effect of online assessment (i.e. whether reviews in an examination collection are overwhelmingly positive or negative) on different

dependent variables, in particular on the perceived utility of the reviews and attitudes towards a product [76–78]. This operationalization of review effectiveness warrants helpfulness and facilitates user's decision-making [79] whilst inspiring trust and confidence among other potential users [80–83]. Nevertheless, the existence of a contradictory total rating reduces the credibility of the analysis due to its negative impact on consumer product attributes [84].

Some scholars are sceptical about the review 'helpfulness' level related to user votes, and suggest other diverse features, such as basic, stylistic, and semantic characteristics by applying data mining techniques from online user reviews [40]. Although this multi-process conceptualization of multiple determinants for review effectiveness is emerging in the literature, some scholars are of the view that participation in reviews and polarity of reviews may be driven by different consumer behaviors [79] and emotions [85] so should be studied separately. Although it is observed that positive review sentiments increase readership [86].

The basis for digital health research is often primarily based on cataloging the number and types of apps dealing with a specific health issue that assess the theoretical effect of such features on user experience or identifies the behavioral concepts in a subset of mHealth apps. Studies can also explore a variety of multi-dimensional app features that contribute to a better user experience and eventually alter long-term behaviours [1]. Based on these gaps from the mHealth app rating, our study aims to extend current knowledge on how users' perceptions and app reviews can generate new app rating that may be different than the original/existing app rating. This paper also presents a new methodology, an AI-enabled mHealth app rating scale, which takes multidimensional measures to generate app rating. By adopting text mining as a novel methodology to extract various semantic characteristics from reviewers' texts, this paper evaluates user review determinants on mHealth app ratings.

mHealth app quality rating scale

mHealth applications

The mHealth applications are an important part of the forthcoming next-generation Internet of Healthcare Things (IoHT) [87]. This paper categorises the mHealth apps into the following six types.

Wellness management app. The Wellness management app tracks health aspects such as walking, sleeping, weight, body mass index, temperature, blood pressure, oxygen saturation (oxygen), pulse, consumption of food and water, etc. These apps allow users to monitor aspects of their health and lifestyle.

Disease management app. Aging, demographics, lifestyles, and chronic diseases have been adding an increasing amount of pressure on health care systems around the globe. In order to reduce such pressure, disease management software is employed to assist the health-care system, aiding in monitoring signs/symptoms, tracking medications' intake, and other side effects.

Educational app. Educational applications offer disease-focused education and self-management training.

Self-diagnosis and decision-making app. An emergency patient assessment with a symptom management questionnaire is performed via the self-diagnosis and decision-making app.

Digital therapeutics and rehabilitation. Digital therapeutics and rehabilitation may be used to prevent, control, or treat disease.

Behavioral change and medication. Behavioral improvement and medication mHealth apps identify behavioral changes that improve patients health and facilitate personalized medicine (see Fig 2).

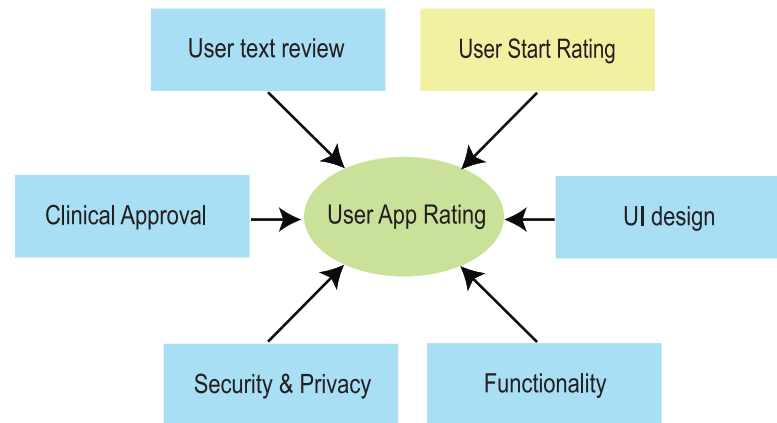


Fig 2. Factors affecting mobile app review.

<https://doi.org/10.1371/journal.pone.0258050.g002>

Factors affecting mHealth app rating

In determining mHealth app quality scores important factors related to the app such as the clinical approval, UI design, functionality, security and privacy are often ignored by the current practice. The complete list of factors which are considered essential in determining an appropriate rating of the mHealth app are provided below:

Approval and certification. Apps, intended for disease diagnosis, cure, mitigation, treatment, or prevention of disease, can be considered as medical devices under the Food and Drug Administration (FDA) regulation [88]. Such an app also has the ability to affect the structure or any function of the human body.

Under the latest guidance of European Medical Device Regulation (EU MDR) [89], apps, having a medical purpose such as prevention or diagnosis of a disease, treatment or alleviation; compensation for an injury or disability; investigation, replacement or modification of anatomy or physiological processes; control of conception; and prediction and prognosis require regulatory approval.

Medical approval is not mandatory for all mHealth applications, as many applications do not engage in clinical diagnosis or intervention [90]. For example, the Wellness Management Apps do not require clinical approval [91].

The mHealth app is designed and developed to meet a specific purpose. The developer offers criteria, specifications, and recommendations after planning the app. Standardization ensures that the facilities, protection, reliability, and consistency of a newly developed mHealth app match the objective. A standardization board must approve all mHealth apps as the mHealth apps need to be budget free and should deliver the expected performance. Authorization and certification provide credibility to a mHealth app and help a physician choose an app for the patient.

Functionality. The app designer's first aim is to create an mHealth app with full functionality and a high-quality experience. Having too many features in an app on a small screen can cause distraction, and it may not be easy for the app user to get all the essential information. The selection of key features is, therefore, essential to ensure consistency of the experience. The well-designed mHealth app must provide expert opinions, meet all problem domain criteria, and provide better accuracy, accompanied by the HCL law [92]. The optimal functioning of the mHealth app can involve alert/ warning, connectivity via multiple interfaces (such as NFC, LiFi, WiFi, Bluetooth, etc.) of a mobile, well documented, stable output, customizable,

low-cost/free, family-friendly, easy-to-use, recording for viewing recorded data. The easy-to-use measure refers to either a qualitative/quantitative assessment of how quickly targeted users use the mHealth app.

UI design. The design of the user interface (UI) is one of the key concerns of the app developer and depends on hardware and technical improvements (such as touch screen / appearance), aesthetics, user context (such as history, age and disability) and functionality. Some of the challenges are limited screen size and a high demand for visual attention. The app designer therefore aims, to provide a sleek interface with important features to ensure the quality of experience.

Security and privacy. Security and privacy are other critical concerns for mHealth. Privacy protects a patient's name, age, sex, cultural, mental, physical/physiological, socioeconomic, and medical factors. A user permission is also required once the third-party agency receives the personal information transfer app, and any unnecessary data must be deleted. In order to maintain privacy, security, and confidentiality, the app developer must choose the appropriate access control and encryption protocol using fewer mobile device resources.

User star rating. User star ratings, also called user star rankings, are average values based on user input on Apps and range from 0 to 5. They are generally averaged over a period of time or by the number of times the App has been updated.

User review. User text reviews provide insight into current trends and problems encountered when using an app. Because of the personalized feedback, the app developer will be able to construct apps based on the user experience and expectation. Thus user text reviews is an important factor to be considered and the normalized value range from 0 to 5.

Methods: AI-enabled text data mining

ACCU³RATE app rating model

The proposed app rating tool, called ACCU³RATE for mHealth Apps, provides a multidimensional application quality evaluation using criteria such as user star rating, user text review, UI design, functionality, security and privacy, and clinical approval. The app rating tool eliminates the aforementioned variables from the app developer and generates an improved rating based on the user experience. In the conventional rating, the quality of the app is judged only by user stars. However, our research suggests that the declaration of the developer as well as user text review must be incorporated. The relationship between input parameters and the model parameters is shown in Fig 3.

Developers declaration

An app developer declares the mHealth app criteria such as clinical approval, standardization certification, functionality, UI design, security and privacy. The app developer's response calculates the initial rating (initial $\in [0, 5]$). Algorithm 1 reflects the complete procedure from calculating the initial rating to changing the rating based on user feedback. At the time of the publication of Apps, the initial rating for an app is created. Based on the user experience (star rating and review), this rating changes. The feature "users recalculation" returns a positive or negative value depending on the polarity of user comments in the Algorithm 1.

Algorithm 1: Initial rating based on developer's feedback

```

Result:  $R_{Base}$ 
total_topic,  $R_{Base} = 0$ ,  $R[total\_topic]$ , counter[total_topic]
if FDA_Approval_Req_Dev_Response() == True then
    total_topic = 5;
else

```

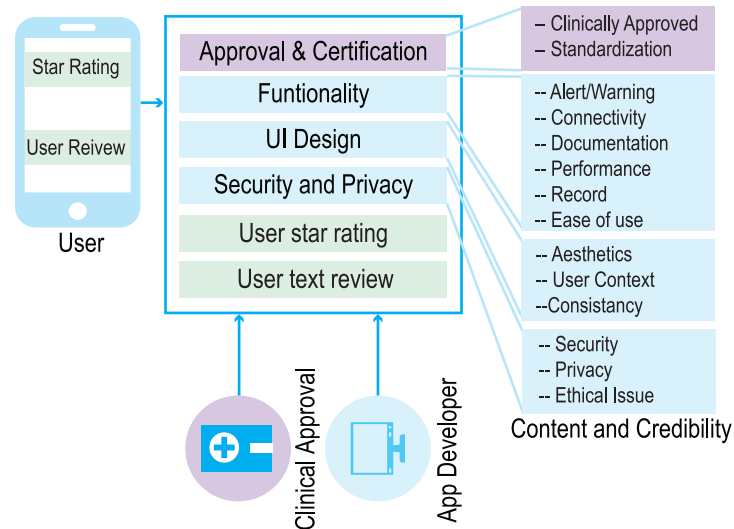


Fig 3. Relationship between input parameters and model parameters.

<https://doi.org/10.1371/journal.pone.0258050.g003>

```

total_topic = 4;
while r = 1 to total_topic do
    R[r] = developer_ans_on_topic_r(); // Initial Rating based on
developer response, 0 or (5/total_topic)
    R_Base += R[r];
end
Publish (R_Base);
    
```

User rating and review

The key variables considered in this paper are discussed as follows.

Star Rating. The app’s star rating is the base goodness metric, which is determined by the cumulative user input rating ranging from 0 to 5. The user considers it as being trustworthy over personal recommendations while downloading an app.

Text mining and sentiment analysis. The polarity of user reviews in the mHealth app can be extracted using the following steps (see Fig 4).

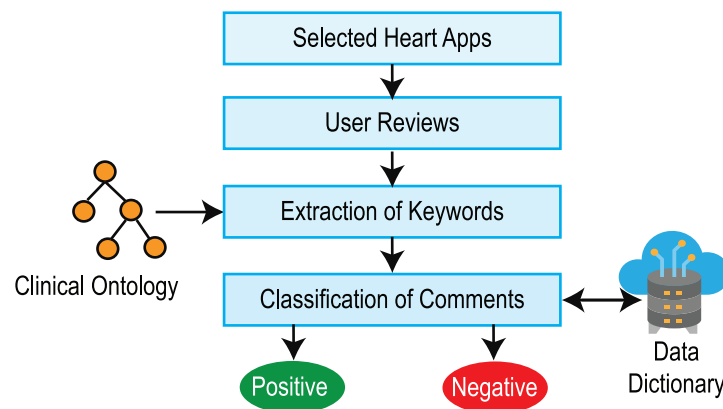


Fig 4. Flow diagram for opinion mining.

<https://doi.org/10.1371/journal.pone.0258050.g004>

Step 1: Selection of mHealth apps: The mHealth applications are selected from the Google Play and App Store for a specific category (such as health-related apps).

Step 2: Collection of comments: Based on the assumption that the user star rating, user review, number of downloaded features, and app features can influence the app's reputation, the aforementioned data for the selected apps is accumulated. NLTK (Natural Language Toolkit) [93], a versatile package of natural language processing algorithms, analyzes the user's critique/comments. All comments from the app review segment are gathered using scrubbing software.

Step 3: Extraction of Keywords: Keywords are extracted from a customer's review that contributes to the app's credibility. We break down the analysis of comments into small bits like words and use tokenization techniques, where word tokenization splits the entire text into words. The frequency that relates to the number of each word in the method is then transmitted and counted. In text data, mining stopwords (i.e., am, are, this, a, an, etc.) is viewed as noise. To delete all of these, we created a stopword list and filtered the tokenized word list with it. After tokenization, the words go through stemming, i.e. a normalization process where the term is shortened to the root word or a derived affix is trimmed.

Step 4: Identification of polarity: The polarity of the comments is defined based on sentiment analysis [94]. We use the clinical ontology and lexicon-based approach to evaluate emotions from a written text to count positive and negative terms among the review comments. The classification of comments is one of the most important aspects in the text mining industry. We created a data dictionary to evaluate our dataset and, with the aid of the **word2vec** algorithm [95], and then we fit the data dictionary with the comment data. Using the NLTK text corpora and **word2vec** algorithm, these review comments are analyzed to find any constructive negative review comments.

We then use the Naive Bayes classifier [96, 97] to evaluate the positive and negative polarities of the comments. We use two environments, NLTK [98] and Spider [99] for sentiment analysis. Word charts in Fig 5 shows the wordcloud for positive and negative polarities of words found in the comments of the app.

Approval and certification

mHealth technologies originated mainly from non-medical practice. As a consequence, such applications can lead to poor content, poor treatment and unfavorable outcomes. Inappropriate use of apps by patients and clinicians could lead to a worsening of health conditions if there is no evidence-based guidelines for app growth. Thus, clinical approval is necessary for mHealth applications especially if the app is used for monitoring of clinical data, provide diagnosis, or intervention for a medical condition.

Strict protocols and criteria are essential for ensuring the quality of an app. The mHealth software developer should have an understanding of the multidisciplinary standards such as ISO 14971:2007 [100] and ISO 14971:2019 [101]. Any mHealth app with clinical approval and ISO certification improves the credibility of the app.

Fig 6 shows calculation of "Clinical Approval" value using four questionnaires such as

- **Q1:** Is the app monitored or analyzed patient data or patient specific medical device data/function?

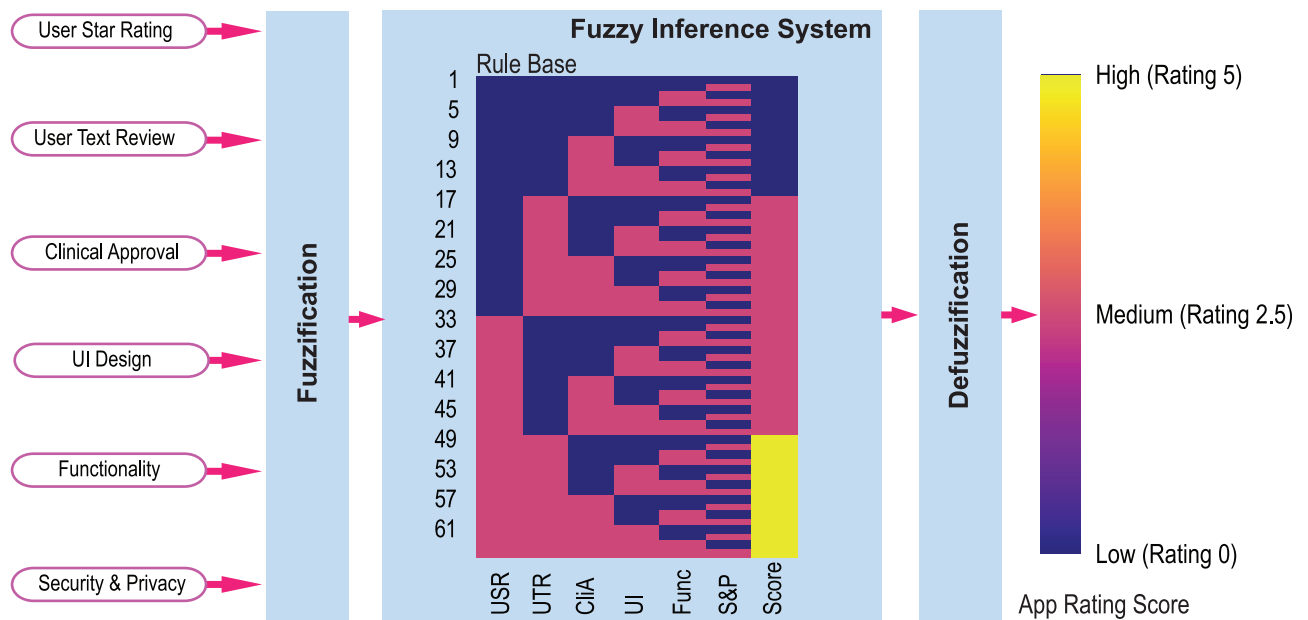


Fig 7. Fuzzy logic based fusion technique which combines the knowledge extracted from the users' star rating, users' text review, clinical approval, UI design, functionality and security & privacy and thereby generates a score.

<https://doi.org/10.1371/journal.pone.0258050.g007>

functionality, security and privacy, and thereby generates an App rating. The FLC has mainly three layers; these are fuzzification, rule base and inference system and defuzzification [103].

The fuzzification layer converts the input variables into fuzzy forms using linguistic terms and Gaussian membership functions that convert the crisp inputs into fuzzy inputs. We have chosen three linguistic terms (Low, Medium, and High) for Users' Star Rating (USR); and two linguistic terms (Low and High) for users' text review (UTR), clinical approval (CliA), UI design (UI), functionality (Func), and Security and Privacy (S&P).

The rule layer includes 62 "IF . . . THEN" rules for the system and a fast forward rule selection algorithm is used to select optimal rules for improving the FLC's accuracy and computation. All the rules are represented by a heatmap in Fig 7

After the fuzzy inferencing, all the consequents are integrated by multiplying a weight value. The defuzzification layer converts the fuzzy value into an average value. In this case, the center of the area, the defuzzification method is employed for its simplicity.

Fig 7 also shows the input and output membership functions, and the surf plots refer to the relationship between the input and output variables. Fig 8 shows the input membership and the relationship between the input-output in the rule base system.

Tool and rating correlation

Cronbach alpha is used to measure interrelated items to estimate quality score (i.e., scale reliability) and internal consistency of the scale. On the other hand, the interrater reliability index, called *Cohen's kappa*, determines the extent to which two raters's ratings agree or disagree with each other for qualitative objects.

Results and discussion

Data collection

The data collection phase started with selecting appropriate apps to provide the proof-of-concept. The apps were searched using the keywords *heart disease*, *heart related medical app*,

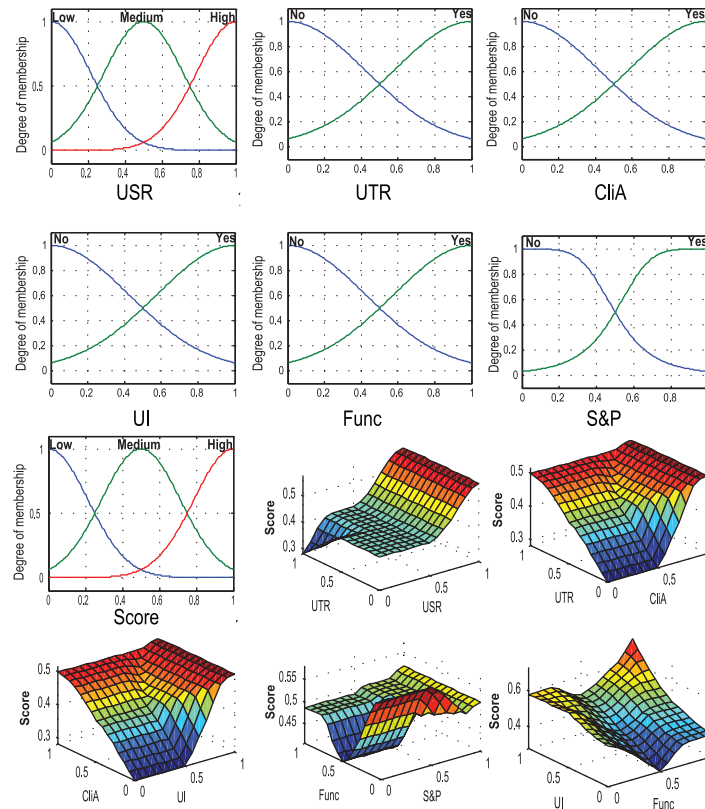


Fig 8. Input membership and relation between the input-output in the rule base system.

<https://doi.org/10.1371/journal.pone.0258050.g008>

healthy heart, and *heart care* in the Google Play Store and Apple store to evaluate the proposed AI-based app rating scale. A total of 317 applications in different categories have been identified based on the search criteria. Each app was initially qualitatively assessed and screened with titles, descriptions, and associated screenshots provided in the Play Store. Out of the total 317 m-Heath apps, 278 are from the Google Play Store while, 39 are from the Apple Store. Then, we set some exclusion criteria for apps, such as duplicate apps, apps not related to heart disease, apps downloaded less than 5000 times, game and simulation apps, apps not updated in more than 12 months, and non-English apps. Consequently, after filtering based on all these criteria, we selected 43 apps (23 apps from Google Play Store and 20 apps from the Apple Store) for further analysis. Fig 9 shows the case study design and app selection flowchart.

Correlation analysis

The proposed mHealth app rating tool provides a multidimensional assessment of the App interaction, functionality, aesthetics, subjective quality, security and privacy.

Fig 10 illustrates the correlation coefficient of selected parameters such as user experience (user star rating and text review), approval and certification, UI design, security and privacy. The correlation coefficient (r) between the variable is considered for the app scale assessment. User experience has found to have influence on functionality ($r = 0.5$), privacy and security ($r = 0.2$). The UI design also affects star rating ($r = 0.1$) and text review ($r = 0.1$). App functionality is correlated with user star rating ($r = 0.5$) and text review ($r = 0.5$). The fuzzy logic based mHealth app evaluation framework has been employed to produce app ratings based on user star rating, user text review, clinical approval, UI design, functionality, and security and

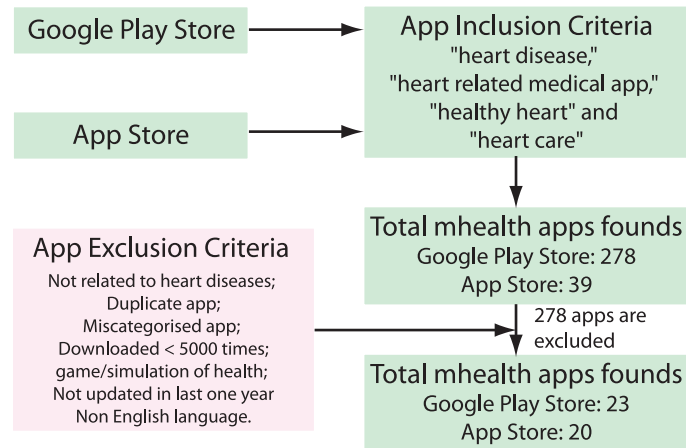


Fig 9. Case study design and app selection flowchart.

<https://doi.org/10.1371/journal.pone.0258050.g009>

privacy. From the initially selected 317 apps, our filtering categories (such as consistency, user context, aesthetics, ethical issues, privacy, security, connectivity, alertness, ease of use, record and clinical approval) informed the final selection of 43 apps. The input factors, proposed in this paper, include various model parameters such as users’ star rating, users’ text review, clinical approval, UI design, functionality, and security and privacy (See Fig 3).

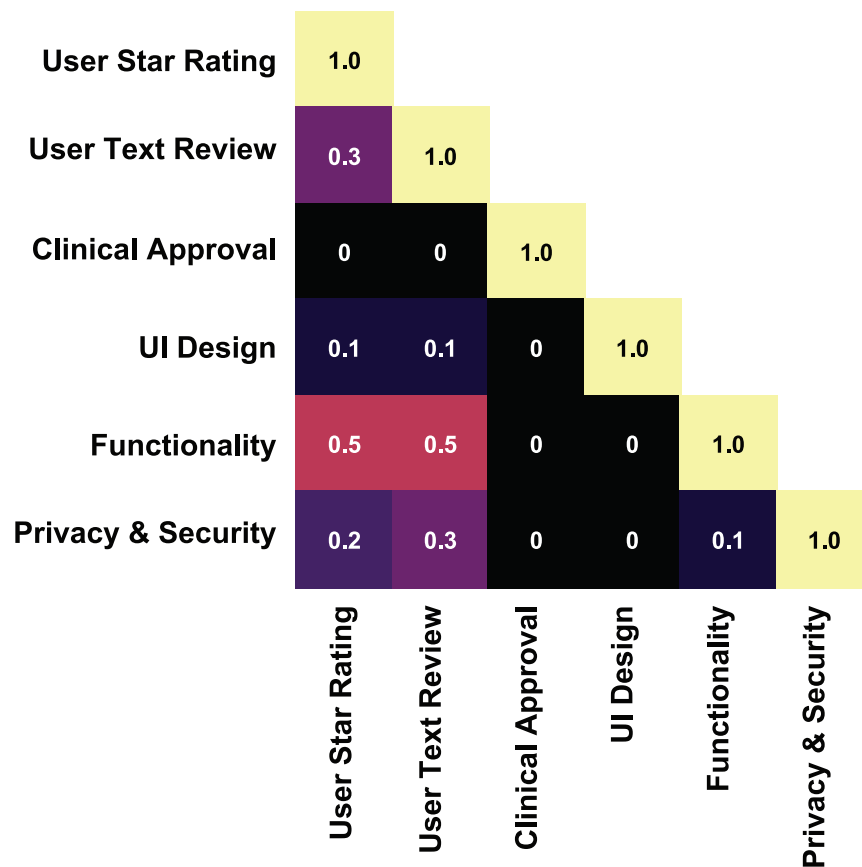


Fig 10. Correlation matrix illustrating the coefficient among factors considered for the assessment of the app scale.

<https://doi.org/10.1371/journal.pone.0258050.g010>

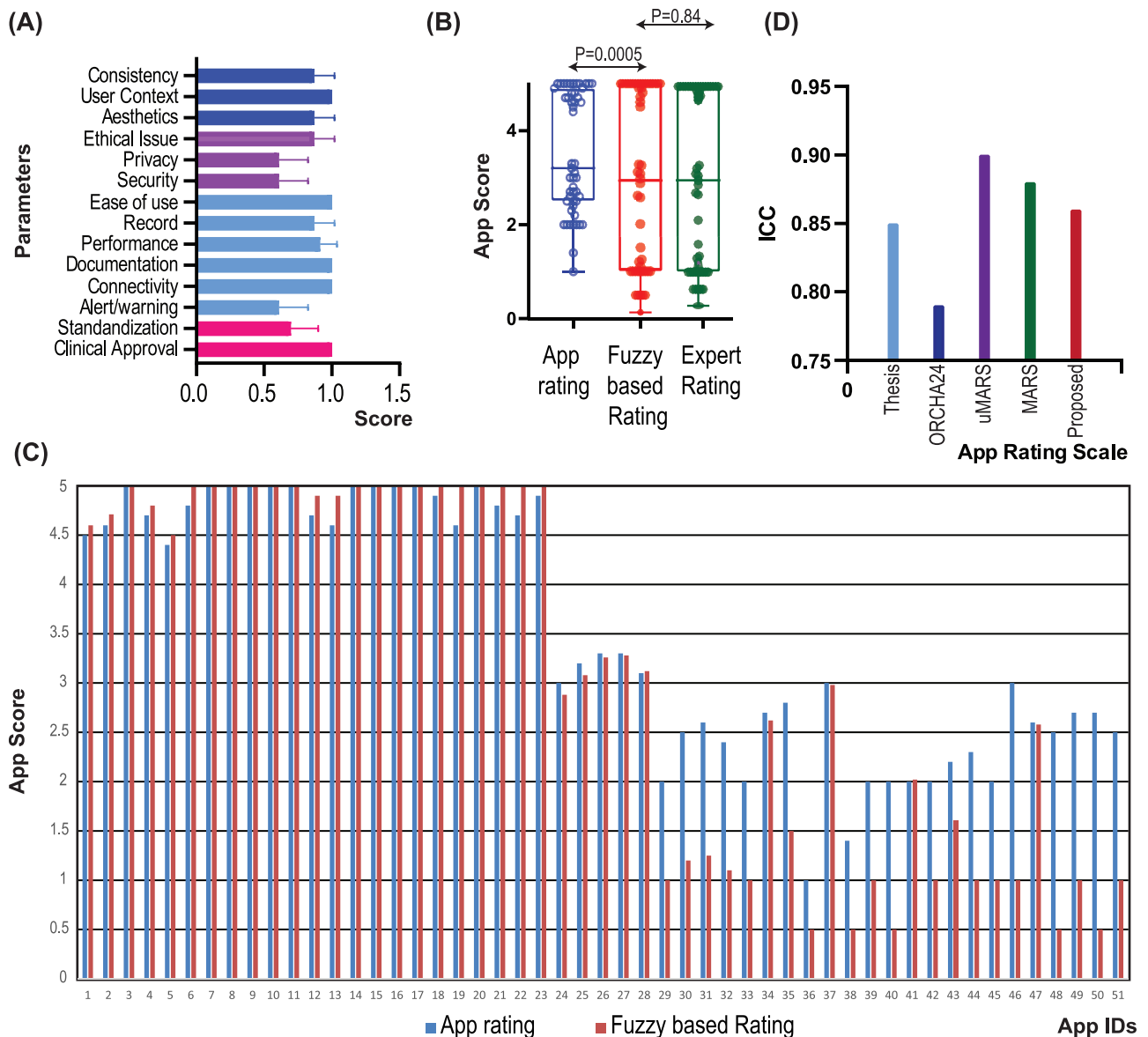


Fig 11. Comparison of traditional mobile app ratings and fuzzy based app ratings. (A) Average scores with respect to each parameter for all the selected 43 apps, (B) Box plot for the average rated value for traditional, fuzzy based ratings, as well as ratings based on expert opinion, (C) app scores for traditional and fuzzy based ratings for all the selected apps, and (D) Comparison of various app scales in terms of ICC.

<https://doi.org/10.1371/journal.pone.0258050.g011>

Fig 11(A) shows the average score for the different parameters, while, Fig 11(B) presents the box plot for the average rated value for conventional app rating scores, fuzzy based scores and app scores based on expert opinion. It has been found that the fuzzy based rating has a high variance compared to the conventional app rating, whereas the fuzzy based rating shows a high relationship in contrast to scoring based on expert opinion ($p = 0.84$). Fig 11(C) demonstrates app scores for conventional and fuzzy-based ratings for all the apps chosen. It has been noted that fuzzy based rating decreases the rating value if the selected input parameters have not been found in these 43 apps.

It has been found that the *Cohen's kappa* value varied as: approval and certification (0.44), functionality (0.41), UI design (0.50), security and privacy (0.55). On the other hand, the

Cronbach alpha value is found to be 0.86 which shows the ACCU³RATE has an acceptable degree of reliability of scale and internal consistency.

When the proposed app rating scale ACCU³RATE is compared with the THESIS, ORCHA-24, uMARS, and MARS in terms of the ICC on 43 apps (23 apps from the Google Play Store and 20 apps from the Apple Store), it is found that the factors in the same domain have 2-way mixed ICC = .86 (95% CI 0.81–0.92) in contrast to the app scales THESIS: 0.85; ORCHA-24: 0.79; uMARS: 0.90; MARS: 0.88 (See Fig 11(D))

The high value of ICC (ICC = .86) shows components under the same criteria in the proposed app scale resemble one another very closely, implying that the ACCU³RATE app scale is very reliable.

Conclusion

This article describes a multidimensional mHealth App rating scale that uses user star ratings, user reviews, and developer-declared features to build an app rating. However, there is currently a dearth of conceptual knowledge on how user reviews effect App ratings on a multidimensional level. This study employs an AI-based text mining technique to gain a more thorough knowledge of user inputs based on a variety of parameters. Six variables were identified as influencing factors for the mHealth app rating scale. These factors are— user star rating, user text review, UI design, functionality, security and privacy, and clinical approval & certification and lays the basis of the proposed mHealth app rating tool named ACCU³RATE. The user sentiment were extracted from their reviews and clinical approval facts, if present, were extracted from the developer statement using the Natural Language Toolkit package. Finally, fuzzy logic was used to combine all of the data and produce the rating score. Several heart-related apps available in the Google Play Store and Apple app store were used as a case research to assess the effectiveness of the app rating scale. The ACCU³RATE mHealth scale has shown excellent reliability as well as internal consistency of the scale, and a high interrater reliability index (Cronbach alpha = 0.86). It has also been found that ratings provided by ACCU³RATE has a high similarity score ($p = 0.84$) to manual scoring performed by experts. The paper has practical implications for App developers who may draw on this knowledge to improve their Apps functionality. Even for mHealth users, new insights are provided on mHealth usability that may help reduce the burdens of healthcare provision through empowering users to manage their own health using innovative mHealth technologies.

Future work

The immediate next step of our research will be to expand the scale, deploying the presented method for a wider range of applications beyond heart disease to verify the generability of the findings. With the better understanding of deep learning (DL) algorithms [97, 104] and increasing demand for explainability and causability in the AI/DL model by clinicians while not compromising on performance, called eXplainable AI (XAI) [105, 106], the proposed model can be extended by considering ML/DL approaches to design a new mHealth apps rating scale which has the Influence of User Reviews.

Supporting information

S1 Data.
(TXT)

Author Contributions

Conceptualization: M. Shamim Kaiser, Atika Ahmad Kemal.

Data curation: M. Shamim Kaiser, Russell Kabir, Mufti Mahmud, Atika Ahmad Kemal.

Formal analysis: Milon Biswas, M. Shamim Kaiser, Russell Kabir, Atika Ahmad Kemal.

Investigation: Marzia Hoque Tania, Mufti Mahmud.

Methodology: M. Shamim Kaiser, Mufti Mahmud.

Project administration: Marzia Hoque Tania.

Supervision: M. Shamim Kaiser.

Writing – original draft: Milon Biswas, Marzia Hoque Tania, M. Shamim Kaiser, Russell Kabir, Atika Ahmad Kemal.

Writing – review & editing: M. Shamim Kaiser, Russell Kabir, Mufti Mahmud, Atika Ahmad Kemal.

References

1. Mendiola MF, Kalnicki M, Lindenauer S. Valuable features in mobile health apps for patients and consumers: content analysis of apps and user ratings. *JMIR mHealth and uHealth*. 2015; 3(2):e40. <https://doi.org/10.2196/mhealth.4283> PMID: 25972309
2. Lee K, Kwon H, Lee B, Lee G, Lee JH, Park YR, et al. Effect of self-monitoring on long-term patient engagement with mobile health applications. *PLoS one*. 2018; 13(7):e0201166. <https://doi.org/10.1371/journal.pone.0201166> PMID: 30048546
3. Kumar S, Viral R, Deep V, Sharma P, Kumar M, Mahmud M, et al. Forecasting major impacts of COVID-19 pandemic on country-driven sectors: challenges, lessons, and future roadmap. *Personal and Ubiquitous Computing*. 2021; p. 1–24. <https://doi.org/10.1007/s00779-021-01530-7> PMID: 33815032
4. Kaiser MS, Mahmud M, Noor MBT, Zenia NZ, Al Mamun S, Mahmud KA, et al. iWorkSafe: towards healthy workplaces during COVID-19 with an intelligent pHealth App for industrial settings. *IEEE Access*. 2021; 9:13814–13828. <https://doi.org/10.1109/ACCESS.2021.3050193>
5. Frost D, Mahmud M. Strengthening Health Systems in Low-Income Countries: A Stakeholder Engagement Framework. *EGOV-CeDEM-ePart 2020*. 2020; p. 215.
6. Kaiser MS, Lwin KT, Mahmud M, Hajjalizadeh D, Chaipimonplin T, Sarhan A, et al. Advances in crowd analysis for urban applications through urban event detection. *IEEE Transactions on Intelligent Transportation Systems*. 2017; 19(10):3092–3112. <https://doi.org/10.1109/TITS.2017.2771746>
7. Kaiser MS, Zenia N, Tabassum F, Al Mamun S, Rahman MA, Islam MS, et al. 6G Access Network for Intelligent Internet of Healthcare Things: Opportunity, Challenges, and Research Directions. In: *Proceedings of International Conference on Trends in Computational and Cognitive Engineering*. Springer; 2021. p. 317–328.
8. Free C, Phillips G, Galli L, Watson L, Felix L, Edwards P, et al. The effectiveness of mobile-health technology-based health behaviour change or disease management interventions for health care consumers: a systematic review. *PLoS med*. 2013; 10(1):e1001362. <https://doi.org/10.1371/journal.pmed.1001362> PMID: 23349621
9. Fiordelli M, Diviani N, Schulz PJ. Mapping mHealth research: a decade of evolution. *Journal of medical Internet research*. 2013; 15(5):e95. <https://doi.org/10.2196/jmir.2430> PMID: 23697600
10. Biduski D, Bellei EA, Rodriguez JPM, Zaina LAM, De Marchi ACB. Assessing long-term user experience on a mobile health application through an in-app embedded conversation-based questionnaire. *Computers in Human Behavior*. 2020; 104:106169. <https://doi.org/10.1016/j.chb.2019.106169>
11. Asif-Ur-Rahman M, Afsana F, Mahmud M, Kaiser MS, Ahmed MR, Kaiwartya O, et al. Toward a heterogeneous mist, fog, and cloud-based framework for the internet of healthcare things. *IEEE Internet of Things Journal*. 2018; 6(3):4049–4062. <https://doi.org/10.1109/JIOT.2018.2876088>
12. Farhin F, Kaiser MS, Mahmud M. Towards Secured Service Provisioning for the Internet of Healthcare Things. In: *2020 IEEE 14th International Conference on Application of Information and Communication Technologies (AICT)*. IEEE; 2020. p. 1–6.

13. Esha NH, Tasmim MR, Huq S, Mahmud M, Kaiser MS. Trust IoHT: A Trust Management Model for Internet of Healthcare Things. In: Proceedings of International Conference on Data Science and Applications. Springer; 2021. p. 47–57.
14. Mukherjee H, Sreerama P, Dhar A, Obaidullah SM, Roy K, Mahmud M, et al. Automatic Lung Health Screening Using Respiratory Sounds. *Journal of Medical Systems*. 2021; 45(2):1–9. <https://doi.org/10.1007/s10916-020-01681-9> PMID: 33426615
15. Sumi AI, Zohora MF, Mahjabeen M, Faria TJ, Mahmud M, Kaiser MS. fASSERT: A fuzzy assistive system for children with autism using internet of things. In: International Conference on Brain Informatics. Springer; 2018. p. 403–412.
16. Jesmin S, Kaiser MS, Mahmud M. Artificial and internet of healthcare things based Alzheimer care during COVID 19. In: International Conference on Brain Informatics. Springer; 2020. p. 263–274.
17. Farhin F, Kaiser MS, Mahmud M. Secured Smart Healthcare System: Blockchain and Bayesian Inference Based Approach. In: Proceedings of International Conference on Trends in Computational and Cognitive Engineering. Springer; 2021. p. 455–465.
18. Faria TH, Shamim Kaiser M, Hossian CA, Mahmud M, Al Mamun S, Chakraborty C. Smart City Technologies for Next Generation Healthcare. In: Chakraborty C, Lin JCW, Alazab M, editors. Data-Driven Mining, Learning and Analytics for Secured Smart Cities: Trends and Advances. Cham: Springer International Publishing; 2021. p. 253–274.
19. Bellei EA, Biduski D, Cechetti NP, De Marchi ACB. Diabetes Mellitus m-Health Applications: A Systematic Review of Features and Fundamentals. *Telemedicine and e-Health*. 2018; 24(11):839–852. <https://doi.org/10.1089/tmj.2017.0230> PMID: 29470105
20. Wang S, Lee HS, Choi W. Assessing the informativeness of user reviews on mobile health applications for chronic diseases. *Proceedings of the Association for Information Science and Technology*. 2019; 56(1):790–791. <https://doi.org/10.1002/pra2.178>
21. Debon R, Coleone JD, Bellei EA, De Marchi ACB. Mobile health applications for chronic diseases: A systematic review of features for lifestyle improvement. *Diabetes & Metabolic Syndrome: Clinical Research & Reviews*. 2019; 13(4):2507–2512. <https://doi.org/10.1016/j.dsx.2019.07.016> PMID: 31405669
22. Zohora MF, Tania MH, Kaiser MS, Mahmud M. Forecasting the risk of type ii diabetes using reinforcement learning. In: 2020 Joint 9th International Conference on Informatics, Electronics & Vision (ICIEV) and 2020 4th International Conference on Imaging, Vision & Pattern Recognition (icIVPR). IEEE; 2020. p. 1–6.
23. Krishna S, Boren SA, Balas EA. Healthcare via cell phones: a systematic review. *Telemedicine and e-Health*. 2009; 15(3):231–240. <https://doi.org/10.1089/tmj.2008.0099> PMID: 19382860
24. BOGDANOV V. The Fundamentals of UX Design for mHealth Apps. *The Journal of MHealth*. 2016; 3(1):4–5.
25. Lee E, Han S, Jo SH. Consumer choice of on-demand mHealth app services: Context and contents values using structural equation modeling. *International journal of medical informatics*. 2017; 97:229–238. <https://doi.org/10.1016/j.ijmedinf.2016.10.016> PMID: 27919381
26. Hamine S, Gerth-Guyette E, Faulx D, Green BB, Ginsburg AS. Impact of mHealth chronic disease management on treatment adherence and patient outcomes: a systematic review. *Journal of medical Internet research*. 2015; 17(2):e52. <https://doi.org/10.2196/jmir.3951> PMID: 25803266
27. Choi W, Tulu B. Effective use of user interface and user experience in an mHealth application. In: Proceedings of the 50th Hawaii International Conference on System Sciences; 2017. p. 3803–3812.
28. Brewer AC, Endly DC, Henley J, Amir M, Sampson BP, Moreau JF, et al. Mobile applications in dermatology. *JAMA dermatology*. 2013; 149(11):1300–1304. <https://doi.org/10.1001/jamadermatol.2013.5517> PMID: 24067948
29. Stevens DJ, Jackson JA, Howes N, Morgan J. Obesity surgery smartphone apps: a review. *Obesity surgery*. 2014; 24(1):32–36. <https://doi.org/10.1007/s11695-013-1010-3> PMID: 23749609
30. Anderson K, Burford O, Emmerton L. Mobile Health Apps to Facilitate Self-Care: A Qualitative Study of User Experiences. *PLOS ONE*. 2016; 11(5):1–21. <https://doi.org/10.1371/journal.pone.0156164> PMID: 27214203
31. Cho J. The impact of post-adoption beliefs on the continued use of health apps. *International journal of medical informatics*. 2016; 87:75–83. <https://doi.org/10.1016/j.ijmedinf.2015.12.016> PMID: 26806714
32. Nelson LA, Mulvaney SA, Johnson KB, Osborn CY. mHealth Intervention Elements and User Characteristics Determine Utility: A Mixed-Methods Analysis. *Diabetes Technology & Therapeutics*. 2017; 19(1):9–17. <https://doi.org/10.1089/dia.2016.0294> PMID: 28099052
33. Buildfire. buildfire Mobile App Download and Usage Statistics (2021); 2021. <https://bit.ly/39c4bUs>.

34. Chen Y, Yang L, Zhang M, Yang J. Central or peripheral? Cognition elaboration cues' effect on users' continuance intention of mobile health applications in the developing markets. *International journal of medical informatics*. 2018; 116:33–45. <https://doi.org/10.1016/j.ijmedinf.2018.04.008> PMID: 29887233
35. Krishnan G, Selvam G. Factors influencing the download of mobile health apps: Content review-led regression analysis. *Health Policy and Technology*. 2019; 8(4):356–364. <https://doi.org/10.1016/j.hlpt.2019.09.001>
36. Kim SM, Pantel P, Chklovski T, Pennacchiotti M. Automatically assessing review helpfulness. In: *Proceedings of the 2006 Conference on empirical methods in natural language processing*; 2006. p. 423–430.
37. Watkins J, Fabielli M, Mahmud M. Sense: a student performance quantifier using sentiment analysis. In: *2020 International Joint Conference on Neural Networks (IJCNN)*. IEEE; 2020. p. 1–6.
38. Wei CP, Yang CC, Lin CM. A latent semantic indexing-based approach to multilingual document clustering. *Decision Support Systems*. 2008; 45(3):606–620. <https://doi.org/10.1016/j.dss.2007.05.008>
39. Sidorova A, Evangelopoulos N, Valacich JS, Ramakrishnan T. Uncovering the intellectual core of the information systems discipline. *Mis Quarterly*. 2008; p. 467–482. <https://doi.org/10.2307/25148852>
40. Cao Q, Duan W, Gan Q. Exploring determinants of voting for the “helpfulness” of online user reviews: A text mining approach. *Decision Support Systems*. 2011; 50(2):511–521. <https://doi.org/10.1016/j.dss.2010.11.009>
41. Rabby G, Azad S, Mahmud M, Zamli KZ, Rahman MM. Teket: a tree-based unsupervised keyphrase extraction technique. *Cognitive Computation*. 2020; 12(4):811–833. <https://doi.org/10.1007/s12559-019-09706-3>
42. Levine DM, Co Z, Newmark LP, Groisser AR, Holmgren AJ, Haas JS, et al. Design and testing of a mobile health application rating tool. *NPJ Digital Medicine*. 2020; 3(1):1–7. <https://doi.org/10.1038/s41746-020-0268-9> PMID: 32509971
43. Kao CK, Liebovitz DM. Consumer mobile health apps: current state, barriers, and future directions. *PM&R*. 2017; 9(5):S106–S115. PMID: 28527495
44. Hamel MB, Cortez NG, Cohen IG, Kesselheim AS. FDA regulation of mobile health technologies. *The New England journal of medicine*. 2014; 371(4):372. <https://doi.org/10.1056/NEJMhle1403384>
45. Liu C, Zhu Q, Holroyd KA, Seng EK. Status and trends of mobile-health applications for iOS devices: A developer's perspective. *Journal of Systems and Software*. 2011; 84(11):2022–2033. <https://doi.org/10.1016/j.jss.2011.06.049>
46. Dey AK. Understanding and using context. *Personal and ubiquitous computing*. 2001; 5(1):4–7. <https://doi.org/10.1007/s007790170019>
47. Bhattacharyya S, Saravanagru R, Thangavelu A. Context aware healthcare application. *International Journal of Computer Applications*. 2011; 975:8887.
48. Bricon-Souf N, Newman CR. Context awareness in health care: A review. *international journal of medical informatics*. 2007; 76(1):2–12. <https://doi.org/10.1016/j.ijmedinf.2006.01.003> PMID: 16488663
49. Carroll JK, Moorhead A, Bond R, LeBlanc WG, Petrella RJ, Fiscella K. Who uses mobile phone health apps and does use matter? A secondary data analytics approach. *Journal of medical Internet research*. 2017; 19(4):e125. <https://doi.org/10.2196/jmir.5604> PMID: 28428170
50. Klasnja P, Pratt W. Healthcare in the pocket: mapping the space of mobile-phone health interventions. *Journal of biomedical informatics*. 2012; 45(1):184–198. <https://doi.org/10.1016/j.jbi.2011.08.017> PMID: 21925288
51. Dicianno BE, Parmanto B, Fairman AD, Crytzer TM, Yu DX, Pramana G, et al. Perspectives on the evolution of mobile (mHealth) technologies and application to rehabilitation. *Physical therapy*. 2015; 95(3):397–405. <https://doi.org/10.2522/ptj.20130534> PMID: 24925075
52. Wattanapisit A, Teo CH, Wattanapisit S, Teoh E, Woo WJ, Ng CJ. Can mobile health apps replace GPs? A scoping review of comparisons between mobile apps and GP tasks. *BMC medical informatics and decision making*. 2020; 20(1):5. <https://doi.org/10.1186/s12911-019-1016-4> PMID: 31906985
53. Zhao J, Freeman B, Li M. Can mobile phone apps influence people's health behavior change? An evidence review. *Journal of medical Internet research*. 2016; 18(11):e287. <https://doi.org/10.2196/jmir.5692> PMID: 27806926
54. Elavsky S, Smahel D, Machackova H. Who are mobile app users from healthy lifestyle websites? Analysis of patterns of app use and user characteristics. *Translational behavioral medicine*. 2017; 7(4):891–901. <https://doi.org/10.1007/s13142-017-0525-x> PMID: 28929368
55. Dahlke DV, Fair K, Hong YA, Beaudoin CE, Pulczynski J, Ory MG. Apps seeking theories: results of a study on the use of health behavior change theories in cancer survivorship mobile apps. *JMIR mHealth and uHealth*. 2015; 3(1):e31. <https://doi.org/10.2196/mhealth.3861>

56. Litman L, Rosen Z, Spierer D, Weinberger-Litman S, Goldschein A, Robinson J. Mobile exercise apps and increased leisure time exercise activity: a moderated mediation analysis of the role of self-efficacy and barriers. *Journal of medical Internet research*. 2015; 17(8):e195. <https://doi.org/10.2196/jmir.4142> PMID: 26276227
57. Zhou L, Bao J, Parmanto B. Systematic review protocol to assess the effectiveness of usability questionnaires in mhealth app studies. *JMIR research protocols*. 2017; 6(8):e151. <https://doi.org/10.2196/resprot.7826> PMID: 28765101
58. Byambasuren O, Sanders S, Beller E, Glasziou P. Prescribable mHealth apps identified from an overview of systematic reviews. *NPJ digital medicine*. 2018; 1(1):1–12. <https://doi.org/10.1038/s41746-018-0021-9> PMID: 31304297
59. Dehling T, Gao F, Schneider S, Sunyaev A. Exploring the far side of mobile health: information security and privacy of mobile health apps on iOS and Android. *JMIR mHealth and uHealth*. 2015; 3(1):e8. <https://doi.org/10.2196/mhealth.3672> PMID: 25599627
60. Sunyaev A, Dehling T, Taylor PL, Mandl KD. Availability and quality of mobile health app privacy policies. *Journal of the American Medical Informatics Association*. 2015; 22(e1):e28–e33. <https://doi.org/10.1136/amiajnl-2013-002605> PMID: 25147247
61. Amor-García MÁ, Collado-Borrell R, Escudero-Vilaplana V, Melgarejo-Ortuño A, Herranz-Alonso A, Arija JÁA, et al. Assessing Apps for Patients with Genitourinary Tumors Using the Mobile Application Rating Scale (MARS): Systematic Search in App Stores and Content Analysis. *JMIR mHealth and uHealth*. 2020; 8(7):e17609. <https://doi.org/10.2196/17609> PMID: 32706737
62. Davalbhakta S, Advani S, Kumar S, Agarwal V, Bhojar S, Fedirko E, et al. A systematic review of the smartphone applications available for coronavirus disease 2019 (COVID19) and their assessment using the mobile app rating scale (MARS). *medRxiv*. 2020. <https://doi.org/10.1101/2020.07.02.20144964> PMID: 32637969
63. Van Singer M, Chatton A, Khazaal Y. Quality of smartphone apps related to panic disorder. *Frontiers in psychiatry*. 2015; 6:96. <https://doi.org/10.3389/fpsy.2015.00096> PMID: 26236242
64. Stoyanov SR, Hides L, Kavanagh DJ, Wilson H. Development and validation of the user version of the Mobile Application Rating Scale (uMARS). *JMIR mHealth and uHealth*. 2016; 4(2):e72. <https://doi.org/10.2196/mhealth.5849> PMID: 27287964
65. Leigh S, Ouyang J, Mimmagh C. Effective? Engaging? Secure? Applying the ORCHA-24 framework to evaluate apps for chronic insomnia disorder. *Evidence-based mental health*. 2017; 20(4):e20–e20. <https://doi.org/10.1136/eb-2017-102751> PMID: 28947676
66. 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 and uHealth*. 2015; 3(1):e3422. <https://doi.org/10.2196/mhealth.3422> PMID: 25760773
67. Carlo AD, Ghomi RH, Renn BN, Areán PA. By the numbers: ratings and utilization of behavioral health mobile applications. *NPJ digital medicine*. 2019; 2(1):1–8. <https://doi.org/10.1038/s41746-019-0129-6> PMID: 31304400
68. Lin TT, Bautista JR. Understanding the relationships between mHealth Apps' characteristics, trialability, and mHealth literacy. *Journal of Health Communication*. 2017; 22(4):346–354. <https://doi.org/10.1080/10810730.2017.1296508> PMID: 28323546
69. Biviji R, Vest JR, Dixon BE, Cullen T, Harle CA. Factors related to user ratings and user downloads of mobile apps for maternal and infant health: Cross-sectional study. *JMIR mHealth and uHealth*. 2020; 8(1):e15663. <https://doi.org/10.2196/15663> PMID: 32012107
70. News I. Reports find that 80% of 'Approved' mHealth Apps have at least Two Major Security Flaws. *The Journal of mHealth*. 2016; 3(1):17.
71. Azad-Khaneghah P, Neubauer N, Miguel Cruz A, Liu L. Mobile health app usability and quality rating scales: a systematic review. *Disability and Rehabilitation: Assistive Technology*. 2020; p. 1–10. PMID: 31910687
72. Janatkah R, Tabari-Khomeiran R, Asadi-Louyeh A, Kazemnejad E. Usability of a Disease Management Mobile Application as Perceived by Patients With Diabetes. *CIN: Computers, Informatics, Nursing*. 2019; 37(8):413–419. PMID: 31394560
73. Singh K, Drouin K, Newmark LP, Lee J, Faxvaag A, Rozenblum R, et al. Many mobile health apps target high-need, high-cost populations, but gaps remain. *Health Affairs*. 2016; 35(12):2310–2318. <https://doi.org/10.1377/hlthaff.2016.0578> PMID: 27920321
74. Huang HY, Bashir M. Users' adoption of mental health apps: examining the impact of information cues. *JMIR mHealth and uHealth*. 2017; 5(6):e83. <https://doi.org/10.2196/mhealth.6827> PMID: 28659256

75. Salazar A, de Sola H, Failde I, Moral-Munoz JA. Measuring the quality of mobile apps for the management of pain: systematic search and evaluation using the mobile app rating scale. *JMIR mHealth and uHealth*. 2018; 6(10):e10718. <https://doi.org/10.2196/10718> PMID: 30361196
76. Li X, Hitt LM. Self-selection and information role of online product reviews. *Information Systems Research*. 2008; 19(4):456–474. <https://doi.org/10.1287/isre.1070.0154>
77. Purnawirawan N, Eisend M, De Pelsmacker P, Dens N. A meta-analytic investigation of the role of valence in online reviews. *Journal of Interactive Marketing*. 2015; 31:17–27. <https://doi.org/10.1016/j.intmar.2015.05.001>
78. Huang AH, Chen K, Yen DC, Tran TP. A study of factors that contribute to online review helpfulness. *Computers in Human Behavior*. 2015; 48:17–27. <https://doi.org/10.1016/j.chb.2015.01.010>
79. Zhang KZ, Zhao SJ, Cheung CM, Lee MK. Examining the influence of online reviews on consumers' decision-making: A heuristic–systematic model. *Decision Support Systems*. 2014; 67:78–89. <https://doi.org/10.1016/j.dss.2014.08.005>
80. Casalo LV, Flavián C, Guinalú M. Understanding the intention to follow the advice obtained in an online travel community. *Computers in Human Behavior*. 2011; 27(2):622–633. <https://doi.org/10.1016/j.chb.2010.04.013>
81. Willemsen LM, Neijens PC, Bronner F, De Ridder JA. “Highly recommended!” The content characteristics and perceived usefulness of online consumer reviews. *Journal of Computer-Mediated Communication*. 2011; 17(1):19–38. <https://doi.org/10.1111/j.1083-6101.2011.01551.x>
82. Flanagin AJ, Metzger MJ. Trusting expert-versus user-generated ratings online: The role of information volume, valence, and consumer characteristics. *Computers in Human Behavior*. 2013; 29(4):1626–1634. <https://doi.org/10.1016/j.chb.2013.02.001>
83. Mackiewicz J, Yeats D. Product review users' perceptions of review quality: The role of credibility, informativeness, and readability. *IEEE Transactions on Professional Communication*. 2014; 57(4):309–324. <https://doi.org/10.1109/TPC.2014.2373891>
84. Qiu L, Pang J, Lim KH. Effects of conflicting aggregated rating on eWOM review credibility and diagnosticity: The moderating role of review valence. *Decision Support Systems*. 2012; 54(1):631–643. <https://doi.org/10.1016/j.dss.2012.08.020>
85. Xu X. Examining the role of emotion in online consumer reviews of various attributes in the surprise box shopping model. *Decision Support Systems*. 2020; 136:113344. <https://doi.org/10.1016/j.dss.2020.113344>
86. Salehan M, Kim DJ. Predicting the performance of online consumer reviews: A sentiment mining approach to big data analytics. *Decision Support Systems*. 2016; 81:30–40. <https://doi.org/10.1016/j.dss.2015.10.006>
87. Powell AC, Torous J, Chan S, Raynor GS, Shwartz E, Shanahan M, et al. Interrater reliability of mHealth app rating measures: analysis of top depression and smoking cessation apps. *JMIR mHealth and uHealth*. 2016; 4(1):e15. <https://doi.org/10.2196/mhealth.5176> PMID: 26863986
88. Kramer DB, Xu S, Kesselheim AS. Regulation of medical devices in the United States and European Union. In: *The Ethical Challenges of Emerging Medical Technologies*. Taylor and Francis; 2020. p. 41–49.
89. Wilkinson B, van Boxtel R. The medical device regulation of the European Union intensifies focus on clinical benefits of devices. *Therapeutic Innovation & Regulatory Science*. 2020; p. 1–5.
90. Cafazzo JA, Casselman M, Hamming N, Katzman DK, Palmert MR. Design of an mHealth app for the self-management of adolescent type 1 diabetes: a pilot study. *Journal of medical Internet research*. 2012; 14(3):e70. <https://doi.org/10.2196/jmir.2058> PMID: 22564332
91. de la Torre Díez I, García-Zapirain B, López-Coronado M, Rodríguez JJ, del Pozo Vegas C. A new mHealth app for monitoring and awareness of healthy eating: development and user evaluation by Spanish users. *Journal of medical systems*. 2017; 41(7):109. <https://doi.org/10.1007/s10916-017-0753-0> PMID: 28555352
92. Alshora DH, Alsaif S, Ibrahim MA, Ezzeldin E, Almeanazel OT, Abou El Ela AES, et al. Co-stabilization of pioglitazone HCL nanoparticles prepared by planetary ball milling: in-vitro and in-vivo evaluation. *Pharmaceutical Development and Technology*. 2020; p. 1–10. PMID: 32174213
93. Millstein F. Natural language processing with python: natural language processing using NLTK. Frank Millstein; 2020.
94. Bakshi RK, Kaur N, Kaur R, Kaur G. Opinion mining and sentiment analysis. In: *2016 3rd International Conference on Computing for Sustainable Global Development (INDIACom)*. IEEE; 2016. p. 452–455.
95. Lin S, Zhang R, Yu Z, Zhang N, et al.; IOP Publishing. Sentiment Analysis Of Movie Reviews Based On Improved Word2vec And Ensemble Learning. *Journal of Physics: Conference Series*. 2020; 1693(1):012088.

96. Rish I. An empirical study of the naive Bayes classifier. *IJCAI 2001 workshop on empirical methods in artificial intelligence*. 2001; 3(22):41–46.
97. Mahmud M, Kaiser MS, McGinnity TM, Hussain A. Deep learning in mining biological data. *Cognitive Computation*. 2021; 13(1):1–33. <https://doi.org/10.1007/s12559-020-09773-x> PMID: 33425045
98. Loper E, Bird S. NLTK: the natural language toolkit. *arXiv preprint cs/0205028*. 2002.
99. Shekhawat SS, Shringi S, Sharma H. Twitter sentiment analysis using hybrid Spider Monkey optimization method. *Evolutionary Intelligence*. 2020; p. 1–10.
100. Coronato A, Cuzzocrea A. An Innovative Risk Assessment Methodology for Medical Information Systems. *IEEE Transactions on Knowledge and Data Engineering*. 2020. <https://doi.org/10.1109/TKDE.2020.3023553>
101. Thiagarajan H, Larson B, Hatcliff J, Zhang Y. Model-based risk analysis for an open-source PCA pump using AADL error modeling. In: *International Symposium on Model-Based Safety and Assessment*. Springer; 2020. p. 34–50.
102. Mahmud M, Kaiser MS, Rahman MM, Rahman MA, Shabut A, Al-Mamun S, et al. A brain-inspired trust management model to assure security in a cloud based IoT framework for neuroscience applications. *Cognitive Computation*. 2018; 10(5):864–873. <https://doi.org/10.1007/s12559-018-9543-3>
103. Kaiser MS, Chaudary MH, Shah RA, Ahmed KM. Neuro-Fuzzy (NF) based relay selection and resource allocation for cooperative networks. In: *ECTI-CON2010: The 2010 ECTI International Conference on Electrical Engineering/Electronics, Computer, Telecommunications and Information Technology*. IEEE; 2010. p. 244–248.
104. Mahmud M, Kaiser MS, Hussain A, Vassanelli S. Applications of deep learning and reinforcement learning to biological data. *IEEE transactions on neural networks and learning systems*. 2018; 29(6):2063–2079. <https://doi.org/10.1109/TNNLS.2018.2790388> PMID: 29771663
105. Ieracitano C, Mammone N, Hussain A, Morabito FC. A novel explainable machine learning approach for EEG-based brain-computer interface systems. *Neural Computing and Applications*. 2021; p. 1–14.
106. Holzinger A, Malle B, Saranti A, Pfeifer B. Towards multi-modal causability with Graph Neural Networks enabling information fusion for explainable AI. *Information Fusion*. 2021; 71:28–37. <https://doi.org/10.1016/j.inffus.2021.01.008>