Medical Apps – The Road To Trust

Urs-Vito Albrecht[1], **Oliver Pramann**[2], **Ute von Jan**[1]

[1]Peter L. Reichertz Institute for Medical Informatics, University of Braunschweig – Institute of Technology and Hannover Medical School, Hannover, Germany

[2]Kanzlei 34, Rechtsanwälte und Notare, Hannover, Germany

Abstract

Objectives: With health apps and medical apps gaining popularity, assessing their quality still remains an open question. Users are often unaware of potential pitfalls and unsure whether an app can be trusted.
Methods: Following an overview over the current situation (including inherent risks and limits) that also covers existing measures and regulatory questions, additional methods for aiding those preforming app assessments are proposed.
Results: Two methods that may aid various stakeholders in their evaluations are outlined.
Conclusions: The presented tools are currently being evaluated and open for discussion; we believe them to be easy to use.

Keywords

Trust, Peer Review, Certification, Mobile Apps, Mobile Health

Correspondence to:

Urs-Vito Albrecht
Peter L. Reichertz Institute for Medical Informatics, University of Braunschweig – Institute of Technology & Hannover Medical School
Address: Carl-Neuberg-Str. 1, 30625 Hannover, Germany
E-mail: albrecht.urs-vito@mh-hannover.de

EJBI 2015; 11(3):en7–en12
received: June 6, 2014
accepted: September 9, 2014
published: April 20, 2015

1 Introduction

Not least due to their easy usability and the numerous possibilities they offer, mobile smart devices (phones, tablet PCs) and the apps they are running have reached high penetration levels in the general population as well as for professionals. The health sector is no exception: users are confronted with an overwhelming number of apps to chose from.

Unfortunately, independent of where these technologies are used, users are often ignorant about potential risks that they pose. Such risks may for example include misleading or false information (sometimes due to unreliable information sources being used by the authors of the app), but also errors in included algorithms. Shortcomings of the employed hardware, e.g. inadequate sensors integrated in a specific mobile device or differences in the specifications of sensors between devices that appear similar at first glance may also cause problems. In highly sensitive areas such as medicine and health, this is simply unacceptable.

The situation is additionally aggravated by two points. Firstly, users are rarely provided with adequate information that would allow them to assess whether an app is adequate for the desired purpose and can be used without risks for the users themselves or – in the case of professional users – the health of the patients they treat. Problematic apps may also pose a danger to the privacy and integrity of the entrusted health related data. Secondly, existing laws and regulation only apply to a negligible number of apps; depending on the jurisdiction where they are used, such apps usually include those with an intended diagnostic or therapeutic purpose. Thus, currently, users cannot count on app safety being ensured via regulatory measures.

For apps where regulation does either not apply or is not observed (although it would very well be expected), various initiatives of the private sector try to step in and offer certification with the aim of providing a certain level of "security" for users.

Nevertheless, independent of whether regulation applies or not, one shortcoming of all certification processes is that they are often quite time consuming and rather expensive. Thus, these processes are often avoided by developers or providers. In the case of private certification initiatives, users may also often not be aware of their existence. Even if they are, it is often unclear which criteria are applied during the certification process.

As a consequence of the overall situation, users often tend to sift through ratings other users have given an app

in the app stores or they look at information publicized in other places, e.g. blogs and other web resources. Still, information gleaned from such sources is often rather unstructured and incomplete and may also be biased.

Resolving the aforementioned problems will necessitate a collaborative effort by all stakeholders involved, including lawmakers and policy makers (on national as well as international levels) as well as developers, store providers and last but not least, the users themselves.

Definition of Health Apps and Medical Apps

In a relatively recent evaluation performed by Research2guidance in 2013, around 97,000 apps with health related content were counted [1] on various full catalogue app stores and it is estimated that this number continues to grow by approximately 1,000 apps each month. Exact numbers for each of the two categories "health apps" and "medical apps" are hard to come by: the two terms are often used synonymously there is often no clear dividing line between both categories. On the part of the store providers, clear guidelines governing which apps should be assigned to which category are lacking as well. Nevertheless, the differentiation between both types of health related apps is important since both classes have different kinds of inherent risks and limitations [2]; these will be explained later on.

The WHO's definition of health may easily serve as a guide for differentiating between the terms "health" and "medical" and thus also for differentiating between apps for both areas:

"Health is a state of complete physical, mental and social well-being and not merely the absence of disease or infirmity" [3]. Following this definition, "health apps" include mobile applications that aim at influencing the physical, emotional and social well-being in a positive manner. While this definition certainly also encompasses apps with an intended diagnostic or therapeutic purpose, apps following this paradigm cover key areas of medicine and as such, should be assigned to the specific category of "medical apps" instead of the more general category of "health apps".

Limitations and Risks

Usage scenarios for both health apps and medical apps are manifold. Nevertheless, even though they open up many new and exciting possibilities, it should not be forgotten that smart devices and the apps running on them have the potential to cause significant harm to patients. Unfortunately, especially for casual users who are unfamiliar with such questions, possible limitations and risks may not be immediately obvious.

Limitations can for example be due to conceptual weaknesses and poor usability. They may also be caused by deficiencies of the included content, the implementation itself or problems of the devices used for running the apps, such as insufficient memory or computing power for the stated purpose of the app. Inadequacies or incompatibilities of the sensors integrated in a mobile device may also be a cause of problems [2].

Trustworthiness and reliability are other aspects that must be kept in mind: It is not only of interest whether a health app or medical app contains credible and valid content and functions as expected, but also how it handles the acquired (medical) data, which tends to be of a highly sensitive nature. Moreover, many apps tend to be more "talkative" than they should be in a medical context. For example, if an app starts to track its users and possibly combines this information with medical and health related data it has been entrusted with, this is highly disturbing and may easily lead to a breach in patient confidentiality. Altogether, the noted points may seriously impact patient care [4].

There is only a limited number of systematic high-quality studies that have taken a closer look at using apps and smart devices for different purposes in a medical setting (e.g. [5, 6]). Often, these studies have a special emphasis on quality, safety and performance of the investigated apps or the notable lack thereof.

A point noted as contributing to shortcomings in the aforementioned areas is the lack of evidence with respect to the presented content [4]. This can be an additional factor contributing to an app's poor quality, independent of whether it targets professional users or private users. Examples of problems that were noted for apps used in the professional domain include the calculation of medical scores or drug dosages; while the performance of most apps was satisfactory, a few of the evaluated apps made clinically relevant calculation errors [6] with potentially serious or even lethal effects on patients' health. On a side note, some of the problems that were identified in apps performing some sort of calculation were due to poor usability rather than erroneous calculations [7].

Apps intended for private use by patients or healthy persons interested in their health and well-being suffer from somewhat similar issues, although the situation may be additionally aggravated by the fact that laypersons are often even less aware of potential problems than professionals. A prime example for this are apps that promise to diagnose possibly malignant lesions on a patient's skin. In a recent study [8] dealing with four apps promising to evaluate pictures taken by a smart devices' camera with respect to potential malignancies, only a single app gave satisfactory results. This app simply sent the acquired images to a human expert for evaluation. For the other three apps – where the evaluation was performed based on automatic image segmentation algorithms – specificity and sensitivity were low and the results were unacceptable.

While the existence and use of such apps may possibly raise awareness and instigate beneficial skin-monitoring behaviour [9], it may be dangerous if users place too much trust in what such an app tells them: at best, the results may raise a false alarm and trouble a patient unnecessarily. On the other hand, if users rely on an app that tells

them that their lesion is nothing to be concerned about although it is malignant, valuable time may go by until professional medical advice is thought. This is an intolerable situation for diseases such as malignant melanoma where timely treatment is essential.

Although the studies mentioned above are far from giving a complete overview over the subject matter, it is apparent that even seemingly professional apps may have serious weaknesses that can lead to serious consequences for patients' health or the safety and integrity of their personal data. For professional users, namely health care professionals using an app as well as developers and distributors of an app, such problems may result in liability issues with all their consequences.

To alleviate the currently unsatisfying situation, some question whether current rules for regulation are sufficiently applied and whether there is professional involvement during the development of apps that are to be used in medical settings. In contrast to [5], we believe that, while certainly being necessary, current rules and regulations cannot be the sole solution. Available regulatory measures do not cover all apps that are intended for a health related field of application. Currently, many apps fall through the figurative cracks of regulation, necessitating additional tools and measures for protecting all stakeholders, from patients to medical professionals as well as developers and distributors.

The following sections will describe basic aspects of regulation as they pertain to health related apps. They will also list currently available options for such apps outside of regulation, and propose measures we would like to see implemented.

2 Apps and Regulation: A Risk Based Approach

As mentioned above, the differentiation between health apps and medical apps is important considering what happens if a product is labeled as a "medical device": If the intended use of a product (as assigned by the manufacturer) – be it hardware or a software application – is medical (e.g., in the context of diagnostics or therapy), conformity to regulation has to be shown. For the US and the European Union, the legal requirements for medical products are basically comparable, but neither here nor there have laws and regulations been adapted to the specifics of medical apps. Thus, existing laws for medical devices are applied. As for other medical products, there are different classes depending on the inherent risks of the product.

In the US, medical apps are subject to an official approval procedure and manufacturers have to compile and submit an extensive list of documents covering risk management, information about software requirements and architecture, quality management, testing and so on. In September 2013, the FDA published a guidance document with specific provisions on how to deal with medical apps [10]. This document is aimed at both manufacturers and FDA personnel and specifies that mobile medical apps that only pose a minimal risk for users are exempted from regulation ("enforcement discretion"). This is a pragmatic approach that befits the overwhelming number of available medical apps and also tries not to impede innovation. The guidance document is only meant as a guide for applying existing regulation to medical apps and is not legally binding [11].

The rules in the EU do not differ much and require similar documentation to be put together. It depends on the risk class assigned to the app whether this documentation simply has to be kept at hand in case of inquiries or whether it has to be submitted to a notified body. In contrast to the US, there are no exemptions for low-risk apps.

Liability

The differentiation between health apps and medical apps is also important in the context of liability. With a few exceptions, e.g. in the case of Android allowing installation of apps from third party sources, mobile apps are usually distributed via official app stores. These stores usually do not make provisions for the peculiarities of medical products. Nevertheless, both for distribution as well as use of apps in a medical setting, there are are legal pitfalls regarding competition law and liability [12]:

Regarding competition law, if an app's intended use (as assigned by the manufacturer) is medical and it is nevertheless distributed without following the appropriate regulatory procedures, this may be a breach of the law against unfair competition and may cause competitors to take legal measures and to claim damages.

Manufacturers can also be held liable if their product, in this case the app being used, is deficient. As already mentioned in the introduction, causes for such defects are manifold. They can range from errors within the software itself to insufficient documentation or inadequate or missing safety notes. All these sources of error may incur liability.

On the other hand, when medical personnel makes use of medical apps while diagnosing or treating patients, they may also be held liable. Any such procedures must follow common professional standards and liability ensues in case anything goes wrong; it is not sufficient to simply put the blame on the manufacturer of the app.

3 What Measures Are Available?

With development cycles of apps shortening dramatically and the rapid integration of innovative new technologies, it is has become quite difficult for users to keep up with developments and to determine whether a medical or health related app meets their needs and demands for quality. The most popular information sources users turn to, i.e. user comments on the app stores as well as various

web pages, blog posts etc. are often not reliable [13]: In most cases, there is little to no background information about these sources and anything posted on such media must be taken with a grain of salt.

There is still an ongoing discussion about quality standards, how they can be applied, appropriate testing be performed and how users can be informed about the results in a transparent manner.

Private Or Commercial Certification

For apps that are not subject to regulation, one possibility is to obtain certification from one of the many, often private, initiatives although this is not mandatory. Certificates these institutions offer are usually based on tests they carry out themselves, but, since many initiatives keep their evaluation criteria under cover, such certification is not necessarily reliable.

This became apparent at the end of 2013, when Happtique, a mobile health solutions company and subsidiary of the Greater New York Hospital Associations for-profit arm GNYHA Ventures, suspended its mobile health app certification program after serious security issues had been found in apps Happtique had previously certified as secure [14]. In addition to such issues, even for private or commercial initiatives, the overwhelming number of available medical apps prohibits an exhaustive evaluation of all available apps. For users, it is therefore a matter of chance whether evaluation results obtained in such a manner are available for an app they are interested in.

Peer Review

Another potential solution that is often suggested is to implement peer review (by experts in the field) for apps, similar to common procedures used in the scientific domain [15]. Although at first glance, user ratings and evaluations published in the app stores may appear somewhat similar to peer review, one should be careful before relying on such ratings. There is usually little to no information about the background and qualification of those voicing their opinion on the stores [13]: are they really qualified and are the statements biased in any way?

Thus, even in such a sensitive field such as medicine, the final decision on whether to trust an app rests with the user. A well-founded decision is only possible based on adequate information and this information should be provided to users in a comprehensive and comprehensible manner, e.g. in the form of an app synopsis [16, 17], via a structured list covering important points (Table 1).

4 Providing Adequate Information

Regulation as well as peer review and certification by private initiatives usually more or less rely on third parties performing the evaluation or assigning a label. Users do play an active role in this context. In addition, the aforementioned processes do not necessarily include adequate information being made available to users upon which they can decide whether they deem a health related or medical app trustable and usable for their specific purpose or not. To be able to do so, users must be provided with comprehensive and easy to use tools they can use. They also need to be made aware of factors they should consider in order not to overlook any important points.

Table 1: Basic categories and criteria for app evaluation, adapted from [16], where a more detailed version is provided.

Criteria	Content
Imprint	Information about the manufacturer or distributor of an app and his associates Meta data of the app
Rationale	Description of the intended purpose(s), target audience, the setting(s) where the app is to be used, its categorization as a medical or non-medical app
Functionality	Description of the functionalities and features included in the app as well as its restrictions and limits Details about the measures used for ensuring good usability
Validity and reliability	Reliability of contained information Description of quality assurance methods used during development
Data requisitioning and management	Amount and types of data collected and processed by the app
Data protection and privacy	Does the manufacturer adhere to data protection and privacy laws; is regulation observed (depends on the intended purpose); jurisdictions involved
Data transmission and storage	Description of measures taken to protect data entrusted to the app (storage and transmission)

App Synopsis

Extensive and accurate information is essential for all evaluations and who is better suited to carry out the task of providing this information than the manufacturer? To simplify all further processes, both for users as well as others performing an evaluation and to improve the comparability between apps, the information should be provided in a structured and standardized manner, e.g. in the form of an app synopsis as it is described in [16] and [2]. Every additional piece of information, especially if it is presented in the structured manner defined by the app synopsis (Table 1, may significantly aid users in evaluating whether an app meets their needs and can be used in a safe manner, even if they are not really familiar with performing such evaluations.

Since users often tend to look for information directly on the app stores, the app synopsis should be provided there as well as on the manufacturer's homepage. Its aim is not to replace what (private or commercial) certifica-

tion initiatives or regulatory processes have to offer. It is rather meant as an additional measure and to provide all stakeholders with an easy to use source of information in cases where regulation and certification do either not apply or prove inadequate and thus, open questions remain.

Although many of the aspects covered by the synopsis have previously already been dealt with by other projects and initiatives, not all of these target apps. For example, the HONCode (Health On the Net Foundation) code of conduct for medical and health related web pages is a prime example of an initiative with somewhat similar criteria [18]. While its listed criteria can certainly be applied for medical apps and health apps, for apps, additional points need to be considered. In contrast to web pages, users often perceive an app as an integral part of their device and not as an extension that would warrant special caution. Due to this misconception, users may easily be tempted to enter information into an app that they would never willingly disclose on a web page that, when combined with other information (e.g., location based data, personal data) already available on the devices can significantly increase the risk.

Providing Additional Support to Users For Self-Performed Ratings

While providing adequate and comprehensive information, e.g. following the aforementioned app synopsis is already helpful, unfortunately, such information is not always being made available.In cases where third party evaluations or certificates are either unavailable or are not deemed sufficient, users are still left to their own devices for assessing whether they can rightfully place their trust in an app.

As outlined in [19], one possible solution in such cases is to provide users with a tool that can aid them in this assessment process, ideally before they download the app they are interested in. While the app synopsis that is to be used by manufacturers and aims at providing as much information as possible in a clearly structured manner, users want to take a swift decision. Therefore, they need a tool that allows them to sift through and rate the available information as befits their needs. If – by chance – a manufacturer has already provided information following the app synopsis, this is an easy to fulfill task, since the proposed user checklist is based on the app synopsis and thus, users can easily match the items of the synopsis and the checklist. In all other cases, the checklist can still guide users through distilling the necessary information from available information sources, e.g. as provided by the manufacturers themselves, but also from third party sources.

The list consists of 38 questions. These questions are divided into 7 categories and cover aspects that are important in the context of rating the safety and effectiveness of both medical as well as health related apps. While all 38 questions touch on points that are important for obtaining a comprehensive picture of an app, for each category, there are one or more key questions. Table 2 lists the 7 categories and corresponding key questions. Depending on the setting an app is to be used in, even a single unsatisfactory answer to a key questions may very well serve as a show stopper; the final decision rests with the user.

Table 2: Categories and corresponding key questions for the user checklist, adapted from [19]

	Category	Key Question(s)
1.	Status of the app	– Is the app a medical product (certification according to regulation) or is there other certification?
2.	Purpose	– Has the app's purpose been clearly stated?
3.	Included functionality	– Are the descriptions of the included functionality thorough and comprehensible, does the functionality match the actual needs?
4.	Risks and limitations	– Are the app's limitations and risks listed (related to medical aspects, data protection & privacy etc.)
5.	Reliability of the content	– Have the identities of the authors/developers and their qualifications been disclosed? – Are the information sources employed for providing the app's content reliable? – Are there potential conflicts of interest leading to a bias in the provided information/functionality?
6.	Data protection & privacy	– Do users remain in control of their data (what is recorded and when); are they informed about how they can influence the collection process? – Can the app be used without entering sensitive data and is data requested on a voluntary basis? – Have adequate precautions for ensuring data protection and privacy during storage and transmission of data been taken and are users informed about their rights in this context?
7.	Imprint	– Is there sufficient contact information specifying where users can turn if they have any questions?

5 Conclusion

Adequate information remains the cornerstone of all evaluations – independent of whether they are performed by notified bodies, by independent initiatives or the users themselves. In addition to the methods we outlined above, i.e. the app synopsis on the manufacturers' side as well as the checklist that is aimed at users, there are various approaches with similar aims, namely supporting all those dealing with app evaluation in a medical context with sufficient information in order to allow them to make informed decisions.

For example, the one-shot pictorial schema developed by Bonacina et al. [20] provides users with a "user oriented ID card" for apps that can draw their attention to the "risky factors of any medical app". Somewhat similar to both our proposed checklist as well as the synopsis, its elements are divided into 6 attribute categories, but instead of simply checking of each item contained in this checklist (with special attention towards the key questions), the pictorial schema uses a traffic light scheme to denote how well an app scores with respect to each element. Due to its visual nature, this pictorial schema is certainly easy to grasp once it has been compiled. As the authors state [20], the pictorial schema can be compiled by "generically interested citizen, a healthcare provider, a doctor, a nurse, the app manufacturer, a declared cohort of users, a scientific society, a governmental body". Nevertheless, we believe that for those who simply want to quickly determine whether an app they are considering meets their demands, and especially for casual users who are unfamiliar with the process of evaluating medical and health related apps, it may be somewhat tedious to compile. Still, due to its granular nature, it is certainly well suited for professional demands.

Aside from existing laws and regulations for medical and health related apps, there are currently many different initiatives wooing for the attention of all stakeholders, but independent of the approach being applied for evaluating and rating an app, the final decision about an app always rests with the users. They have to assess and decide on their own whether the benefits they anticipate from using an app in a medical context can outweigh its perceived risks [19] or whether they would rather get professional advice before making a decision.

In conclusion, there are currently many interesting developments and it remains to be seen where the "road to trust" will lead: It is important to provide users with apps that are trustworthy and well adapted for the settings they are to be applied in. This is essential in order not to gamble away the trust users place in these exciting technologies and thus hinder future innovations. To smoothly pave this "road to trust" will require a collaborative effort from all sides, including lawmakers as well as developers and last but not least the users themselves.

References

[1] Research2Guidance. Mobile Health market Report 2013-2017; [cited 2014 September 8] 2013. Available from: http://www.research2guidance.com/shop/index.php/-mhealth-report-2.

[2] Albrecht UV, Pramann O, von Jan U. Chapter 7: Synopsis for Health Apps Transparency for Trust and Decision Making. In: Househ M, Boryck E, Kushniruk A, editors. Social Media and Mobile Technologies for Healthcare. IGI Global; 2014.

[3] WHO. Preamble to the Constitution of the World Health Organization as adopted by the International Health Conference, New York, 19–22 June, 1946; signed: 22 July 1946 by the representatives of 61 states (Off. Records of the WHO, no. 2, p. 100); entered into force on 7 April 1948; 1948.

[4] Buijink AWG, Visser BJ, Marshall L. Medical apps for smartphones: lack of evidence undermines quality and safety. Evid Based Med. 2013 Jun;18(3):90–92.

[5] Visvanathan A, Hamilton A, Brady RRW. Smartphone apps in microbiology–is better regulation required? Clin Microbiol Infect. 2012 Jul;18(7):E218–E220. Available from http://dx.-doi.org/10.1111/j.1469-0691.2012.03892.x

[6] Bierbrier R, Lo V, Wu RC. Evaluation of the accuracy of smartphone medical calculation apps. J Med Internet Res. 2014;16(2):e32. Available from: http://dx.doi.org/-10.2196/jmir.3062.

[7] Marshall A. Designing telemedicine apps that health commissioners will adopt. In: 14th Conf. of Open Innovations Association (FRUCT 2013); 2013. p. 63–68. Available from: http://ieeexplore.ieee.org/stamp/stamp.jsp?ar-number=6737946.

[8] Wolf JA, Moreau JF, Akilov O, Patton T, English JC 3rd, Ho J, et al. Diagnostic inaccuracy of smartphone applications for melanoma detection. JAMA Dermatol. 2013 Apr;149(4):422–426.

[9] Hamilton AD, Brady RRW. Medical professional involvement in smartphone 'apps' in dermatology. Br J Dermatol. 2012 Jul;167(1):220–221.

[10] FDA. Mobile Medical Applications - Guidance for Industry and Food and Drug Administration Staff; 2013.

[11] Pramann O, Albrecht UV. Medical-Apps – Im internationalen Vergleich. EHEALTHCOM. 2014 1;.

[12] Pramann O. Mobile Endgeräte und medizinische Software-Applikationen im Krankenhaus. In: Gärtner A, editor. Medizintechnik und Informationstechnologie, MIT - Konzepte, Technologien, Anforderungen, 6. Ergänzungslieferung. TÜV Media GmbH, TÜV Rheinland Group; 2012.

[13] Hill S. Gaming the system: How reliable are app store ratings? 2012 [cited 2014 September 8]. Available from: http://www.digitaltrends.com/computing/gaming-the-system-how-reliable-are-app-store-ratings/

[14] Dolan B. Happtique suspends mobile health app certification program; 2013 [cited 2014 September 8]. Available from: http://mobihealthnews.com/28165/happtique-suspends-mobile-health-app-certification-program/.

[15] JMIR mhealth and uhealth – Apps Peer-Review Launched; 2013 [cited 2014 September 8]. Available from: http://mhealth.jmir.org/announcement/view/67.

[16] Albrecht UV. Transparency of health-apps for trust and decision making. J Med Internet Res. 2013;15(12):e277. Available from: http://dx.doi.org/10.2196/jmir.2981.

[17] PLRI MedAppLab. App Synopsis. 2013 [cited 2014 September 8]. Available from: http://www.app-synopsis.de/

[18] Health On the Net Foundation. The HON Code of Conduct for Medical and Health Web Sites (HONcode). 2013 [cited 2014 September 8]. Available from: http://www.hon.ch/HONcode/Patients/Conduct.html

[19] Albrecht UV, Noll C, von Jan U. App-Synopsis: Self-Assessment on Trust or Distrust of Health-Apps. Stud Health Technol Inform. 2014;202:233-6.

[20] Bonacina S, Marceglia S, Pinciroli F. A Pictorial Schema for a Comprehensive User-oriented Identification of Medical Apps. Methods Inf Med. 2014 May;53(3):208–224.

