Aus dem
Institut für Allgemeinmedizin und interprofessionelle Versorgung des Universitätsklinikums Tübingen

Symptom Checker Apps: From Self-Diagnosis to Patient Empowerment? Results from a multidisciplinary research project in Germany

Inaugural-Dissertation
zur Erlangung des Doktorgrades
der Humanwissenschaften

der Medizinischen Fakultät
der Eberhard Karls Universität
zu Tübingen

vorgelegt von

Wetzel, Anna-Jasmin, geb. Klopfer

2025

Dekan:	Professor Dr. B. Pichler
1. Berichterstatterin:	Professorin Dr. S. Joos
2. Berichterstatterin:	Professorin Dr. U. Cress
3. Berichterstatter:	Professor Dr. H. C. Vollmar
Tag der Disputation:	16.04.2025

This work is licensed under a CC BY 4.0 license.

Table of Contents

List of Figures 4
Abbreviations 5
1. Introduction 6
1.1. Relevance of Symptom Checker Apps 6
1.2. Performance of SCA 7
1.3. Impact of SCA on Health Care Systems 8
1.4. Usage of SCA 10
1.5. Aims and Objectives of the Thesis 11
2. Results 13
2.1. Publication 1 13
2.2. Publication 2 25
2.3. Publication 3 37
2.4. Publication 4 47
3. Discussion 59
3.1. Overview and Impact of SCA Usage 59
3.2. Predictors of SCA Usage 60
3.3. Situational Factors in SCA Adoption 61
3.4. Health Anxiety in the Context of SCA Use 62
3.5. Strengths and Limitations 70
3.6. Conclusions 71
4. Summary 73
5. German summary 75
6. References 77
7. Declaration of Own Contributions 83
8. Publications in Connection with the Thesis 87
9. Acknowledgments 88

List of Figures

Figure 1: Potential negative outcomes of SCA use for individuals with health anxiety 68

Abbreviations

DSM-III	Diagnostic and Statistical Manual of Mental Disorders, Third Edition
DSM-IV	Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition
DSM-V	Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition
GP	General Practitioner/s
HISB	Health Information-Seeking Behavior
ICD-10	International Classification of Diseases, Tenth Revision
ICPC	International Classification of Primary Care
IU	Intolerance of uncertainty
NHS	National Health Service
NIH	National Institutes of Health
SCA	Symptom Checker Application/s
UK	United Kingdom
WI	Whiteley Index

1. Introduction

1.1. Relevance of Symptom Checker Apps

Over the past years, there has been a significant increase in the number of health apps available in app stores, such as the Apple App Store[1] or the Google Play Store[2]. Health apps encompass a broad range of applications designed for health, wellness, and medical purposes. These apps are not always subject to regulation and can serve a diverse array of functions related to well-being and health.

Concurrently, there is mounting pressure on the healthcare system in Germany due to, among other factors, an increase in prevalent chronic conditions in an aging society, like diabetes, intensifying the demand for medical services[3]. This surge in demand coincides with a growing shortage of medical professionals, adding further strain to the healthcare system in the upcoming years[3]. Given the evolving scenario, the potential offered by new technologies appears increasingly promising. Despite being the least trusted source for health information by patients compared to physicians, the internet remains a common choice for consultation considering medical complaints. Research into health information seeking behavior (HISB) reveals that 72% of people in Germany use the internet for health information[4].

Symptom Checker Apps (SCA) are one example of a new kind of freely available health apps. Wallace et al. defined SCA as tools that allow users to enter symptoms and biodata to receive a list of possible diagnoses and clinical triage advice[5]. These tools rank potential diagnoses by likelihood and guide users on the most appropriate next steps, which may include seeking emergency care, consulting a general practitioner (GP), or managing self-care[5]. Therefore, SCA may have the potential to reduce the strain on health care systems, by helping allocate resources more efficiently to patients with medical needs[5]. SCA could possibly act as an initial point of contact for patients with the health care system, guiding them to the most suitable healthcare service. Both public and private companies have promoted SCA as cost-effective solution[5]. Software companies like Ada Health advertise their SCA with the following promises: “*We help people*

understand, manage, and get care for their symptoms with trusted medical expertise in minutes. Healthcare services around the world face urgent challenges and uncertain futures. Millions can't access the care they need. We're here to help"[6]. In order to fulfill these promises, SCA must operate at a level comparable to a consultation with a trained physician[7]. This involves personalizing results to align with the patient's medical history and current treatments, and incorporating complex mechanisms that vary based on the expertise and specific medical field considering users complaints[8].

1.2. Performance of SCA

At the moment, triage performance as well as diagnostic accuracy is insufficient in SCA. One of the first studies investigating the performance of SCA was conducted by Semigran et al. in 2015. Twenty-three SCA were investigated using standardized case vignettes. Across all SCA, the correct diagnosis was listed as most probable in 34% of the vignettes[7]. Considering the top 20 diagnoses, 58% of the SCA listed the correct diagnosis. Triage advice was appropriate in 57% of the vignettes. Of note is that the performance of SCA within non-emergency scenarios was lower the less the urgency was of the scenarios[7]. A follow-up study from 2022, based on the audit study of Semigran et al.[9], showed that triage performance had, on average, not improved over the last five years. Furthermore, the triage capability differed strongly between the different SCA. Triage accuracy was close to that in 2015 with 56%[9]. Further, the SCA were less risk averse in 2020 than in 2015[9]. SCA missed more than 40% of emergencies and did not outperform medical laypeople in their decision[9]. The authors emphasize the importance of two metrices 1. the ability to reliably detect emergencies, as this value reflects SCA's safety and 2. the ability to correctly spot cases where self-care is sufficient, as this reflects the usefulness of the SCA. Considering the reported data of two studies, SCA do not perform well on both metrics[9-11]. A systematic review from 2022 that investigated ten studies on SCA found low diagnostic accuracies ranging from 19 – 38% in different studies[5]. The diagnostic accuracy varied between individual SCA, despite a consistent symptom data input[5]. Triage accuracy was found to be higher than diagnostic accuracy, ranging from 50 – 90%[5]. The review also investigated whether SCA performed better in emergency

than in non-emergency scenarios and found conflicting results[5]. A further study compared the triage capability of SCA to those of medical laypersons and found that overall SCA did not perform superior[12]. However, SCA detected emergencies with a higher reliability than medical laypersons (81% vs 68%)[12]. Both SCA and medical laypersons struggled with self-care cases, and SCA more often wrongly classified them as emergencies compared to laypersons (25% vs 4%)[12].

SCA are often evaluated with case vignettes. A study from 2022 investigated from a test-theoretical view whether clinical vignettes are suitable to evaluate SCA performance[13]. It was found that most SCA changed their performance rank order when adjusting the triage capability for item difficulty[13]. This suggests that reported triage accuracy may commonly be overestimating the capability of SCA, as most studies, the cited study did not account for the fact that SCA tend to selectively appraise easier cases[13]. As case vignettes seem only to offer a first indication of SCAs' performance, further research using different approaches is necessary. A recent 2022 study employed a different approach by using simulated patients and found that SCAs identified the correct main diagnosis in 30% of cases, while emergency physicians achieved a correct diagnosis in 81% of cases[8]. Considering triage accuracy, physicians also outperformed the SCA (96% vs 71%)[8].

To conclude, the performance of SCA is insufficient, diagnosis and triage accuracy are low. SCA perform differently well in (non)-emergency scenarios, depending on the used software and the symptoms that are entered. In general, it cannot be ruled out that SCA are unsafe and put patients at risk as a result of undertriaging nor that their use even puts further strain on health care systems by overtriaging patients.

1.3. Impact of SCA on Health Care Systems

The current state of research on the impact of SCA on healthcare systems reveals a notable gap. Currently, SCA are not uniformly integrated into healthcare systems. Instead, they exist in a fragmented landscape, with numerous applications available across various app stores and browsers. This scattered presence makes it challenging to accurately assess their impact on healthcare systems.

Firstly, it is often unclear whether patients have utilized an SCA tool prior to engaging with healthcare services. This lack of visibility into patient interaction with these tools prior to healthcare visits shows the difficulty of assessing their role in the patient journey within a health care system. Secondly, there is a wide variation in the output and performance of different SCA[5,14]. Since the available SCA differ significantly in their functionality and effectiveness, estimating their overall impact becomes complex. The variability in application performance means that any assessment of their impact might not yield adequate or representative results.

Although the performance of SCA is inadequate, they have already been implemented into the national healthcare system (NHS) in the United Kingdom (UK). The NHS111 symptom checker is implemented on the official NHS111 webpage. The National Institute for Health and Care Research (NIH) in the UK conducted a study on the impact of introducing the NHS 111 online symptom checker. The findings indicated a modest effect on the volume of calls to the NHS 111 telephone service, showing a 1.3% increase in triaged calls[15]. The online service led to more recommendations for emergency and urgent care, ranging from 4.2% to 6.7%[15]. The NHS claims that operating both services in parallel could increase annual costs unless a minimum of 38% of telephone contacts shift to online[15]. The authors of the study concluded that the introduction of the NHS 111 online service suggests a potential for creating new demand[15].

In light of the current status of SCA in Germany, it appears that they have a limited impact. A 2023 study highlighted that only a small portion of the German population (16%) is aware of SCA, with a mere 6.5% actively using them[16]. Despite this, there have been ongoing efforts, both past and present, to integrate SCA into the German healthcare system. A German statutory health insurance implemented a symptom checker in 2019 but had to discontinue their cooperation, since it became known that health data was shared with Facebook and other third-party companies[17]. The German “Kassenärztliche Bundesvereinigung” offers a symptom checker on their homepage[18]. Meanwhile, other stakeholders in the healthcare sector in Germany are also offering symptom checkers. The hospital group “Sana” has implemented a symptom checker on its homepage[19]. In

addition, several health insurance companies implemented symptom checker as part of their supply (e.g. AOK[20], AXA[21]). Symptom checkers may gain in relevance with increasing interest from various stakeholders in health care context. Stakeholder interests in these apps vary widely, from financial savings to hopes for better care. Given the diverse group of stakeholders involved, it is crucial to maintain transparency about the motivations and biases that could influence the app's recommendations. This transparency is especially important because the varied interests could impact the reliability and neutrality of the information the apps provide.

1.4. Usage of SCA

Characteristics of SCA users were investigated in recent research. The median age of SCA users is approx. 40 years old[16,22-25] and therefore slightly younger than the median age of the European population (44.4 years). The study results were derived in different countries, including Europe and America. Willingness to use these tools appears to decline with age[22,24,26,27]. While younger users find SCA more useful, older users are more likely to recommend them[23,27]. Gender distribution among SCA users tends to be predominantly female, with estimates ranging from 62%[16] to 85%[24]. SCA users were described as well-educated[16,24,28]. The latter group is also more inclined to use health apps, linking higher eHealth literacy to a greater likelihood of SCA usage[29]. Another significant factor is the individual's trust level: concerns about incorrect diagnoses and increased anxiety are major issues with SCA usage, indicating that trust significantly impacts user interaction with SCA[30]. Most quantitative data on SCA users come from studies focusing on a single SCA. Despite general perceptions of their usefulness, a lack of awareness about SCA has been identified as a potential hindrance to their wider adoption[16]. Another aspect that increases the likelihood of using SCA is limited access to care[15].

Moreover, insights into the demographics and characteristics of the user group have been identified as key elements in understanding and interpreting the results provided by SCA. Factors such as health anxiety, self-efficacy, acceptance of technology, and education level significantly influence how users interact with SCA[7,16,26,31]. Additionally, trust in the SCAs' accuracy and users' pre-existing

medical knowledge also play a role in how they perceive and act upon the advice or results provided by these applications[32,33].

SCA are most commonly used to better understand the causes of present symptoms in their users, followed by triage advice[28]. The majority (70%) of SCA users reported receiving useful information, while only half of the SCA users reported positive health effects[28]. Approximately 65% of the participants in one study reported following the proposed triage advice[25]. A study from 2018 found that people were more likely to follow the SCA advice to contact their GP during office hours (75%), compared to self-care (67%), out-of-hours care (61%), and wait-and-see instructions (56%). Compliance was higher among male patients, and those who were satisfied with the service of the SCA[25]. Non-compliance was often due to difficulties in communicating symptoms, previous doctor consultations, or preference for one's own judgment[25]. A significant limitation of the study was that it measured intended compliance rather than actual care-seeking behavior. A further study reported that the SCA advice to contact emergency services or visit emergency departments was frequently seen as inappropriate and unnecessary by users[34].

When examining the nature of symptoms and organ system concerns reported, the data reveals a significant prevalence of issues related to reproductive, general, and gastrointestinal systems. Among these, pain emerges as the most commonly reported symptom for SCA use. A recent scoping review underscores the value of SCA in dealing with sensitive or stigmatizing conditions[26]. Particularly for medical problems connected to intimacy, these tools appear to lower barriers, making it easier for individuals to seek care for issues they might otherwise hesitate to address[26]. This aspect of SCA use is crucial in enhancing access to healthcare, especially for conditions that might be stigmatized or resulting in discomfort in SCA users.

1.5. Aims and Objectives of the Thesis

The objectives of the present thesis and its publications were based on the state of scientific knowledge at the time of manuscript preparation. The manuscripts presented in this dissertation were created within the multi-disciplinary research

project “CHECK.APP” funded by the Federal Ministry of Education and Research from 31.03.2020 to 31.12.2023 (Grant number: 01GP1907A).

This dissertation seeks to comprehensively explore the use of SCA of medical laypersons in Germany through a mixed-methods approach. The subsequent paragraph outlines the research questions of the recent dissertation, organized into three main themes.

Overview and Impact of SCA Usage (Publication 2)

- What is the extent of SCA usage in Germany and which applications are most prevalent?
- How do SCA integrate within the HISB framework?
- Is there a correlation between SCA usage and the frequency of consultations of the GP, and patient satisfaction with these medical professionals?
- What are the factors contributing to the non-use of SCA?

Predictors of SCA Usage (Publication 3)

- Which user characteristics predict the use of SCA?
- How do these characteristics influence the adoption and sustained use of SCA?

Situational Factors in SCA Adoption (Publication 4)

- What situational factors influence the adoption of SCA among medical laypeople?
- What aspects of SCA use are most significant to users after they have utilized these tools?

Publication 1 outlines the framework of the research project CHECK.APP in a study protocol; therefore, it does not provide answers to the research themes.

2. Results

2.1. Publication 1

Publication 1, titled “Ethical, Legal, and Social Implications of Symptom Checker Apps in Primary Health Care (CHECK.APP): Protocol for an Interdisciplinary Mixed Methods Study”, provides the framework for the multidisciplinary research project CHECK.APP and incorporates the study components relevant to this dissertation. This protocol was published in JMIR Research Protocols (Impact factor for 2023: 1.7).

Protocol

Ethical, Legal, and Social Implications of Symptom Checker Apps in Primary Health Care (CHECK.APP): Protocol for an Interdisciplinary Mixed Methods Study

Anna-Jasmin Wetzel[1*], MSc; Roland Koch[1*], MD; Christine Preiser[2], MA; Regina Müller[3], MA; Malte Klemmt[4], MA; Robert Ranisch[5], Prof Dr; Hans-Jörg Ehni[3], Prof Dr; Urban Wiesing[3], Prof Dr; Monika A Rieger[2], Prof Dr; Tanja Henking[4], Prof Dr; Stefanie Joos[1], Prof Dr

[1]Institute of General Practice and Interprofessional Care, University Hospital Tübingen, Tübingen, Germany

[2]Institute of Occupational and Social Medicine and Health Services Research, University Hospital Tübingen, Tübingen, Germany

[3]Institute of Ethics and History of Medicine, University Tübingen, Tübingen, Germany

[4]Institute of Applied Social Science, University of Applied Science Würzburg-Schweinfurt, Würzburg, Germany

[5]Faculty of Health Science Brandenburg, University of Potsdam, Potsdam, Germany

[*]these authors contributed equally

Corresponding Author:
Anna-Jasmin Wetzel, MSc
Institute of General Practice and Interprofessional Care
University Hospital Tübingen
Osianderstr. 5
Tübingen, 72076
Germany
Phone: 49 7071 29 85226
Email: jasmin.wetzel@med.uni-tuebingen.de

Abstract

Background: Symptom checker apps (SCAs) are accessible tools that provide early symptom assessment for users. The ethical, legal, and social implications of SCAs and their impact on the patient-physician relationship, the health care providers, and the health care system have sparsely been examined. This study protocol describes an approach to investigate the possible impacts and implications of SCAs on different levels of health care provision. It considers the perspectives of the users, nonusers, general practitioners (GPs), and health care experts.

Objective: We aim to assess a comprehensive overview of the use of SCAs and address problematic issues, if any. The primary outcomes of this study are empirically informed multi-perspective recommendations for different stakeholders on the ethical, legal, and social implications of SCAs.

Methods: Quantitative and qualitative methods will be used in several overlapping and interconnected study phases. In study phase 1, a comprehensive literature review will be conducted to assess the ethical, legal, social, and systemic impacts of SCAs. Study phase 2 comprises a survey that will be analyzed with a logistic regression. It aims to assess the user degree of SCAs in Germany as well as the predictors for SCA usage. Study phase 3 will investigate self-observational diaries and user interviews, which will be analyzed as integrated cases to assess user perspectives, usage pattern, and arising problems. Study phase 4 will comprise GP interviews to assess their experiences, perspectives, self-image, and concepts and will be analyzed with the basic procedure by Kruse. Moreover, interviews with health care experts will be conducted in study phase 3 and will be analyzed by using the reflexive thematical analysis approach of Braun and Clark.

Results: Study phase 1 will be completed in November 2021. We expect the results of study phase 2 in December 2021 and February 2022. In study phase 3, interviews are currently being conducted. The final study endpoint will be in February 2023.

Conclusions: The possible ethical, legal, social, and systemic impacts of a widespread use of SCAs that affect stakeholders and stakeholder groups on different levels of health care will be identified. The proposed methodological approach provides a multifaceted and diverse empirical basis for a broad discussion on these implications.

Trial Registration: German Clinical Trials Register (DRKS) DRKS00022465; https://tinyurl.com/yx53er67

XSL•FO
RenderX

International Registered Report Identifier (IRRID): DERR1-10.2196/34026

(JMIR Res Protoc 2022;11(5):e34026) doi: 10.2196/34026

KEYWORDS

symptom checker apps; self-diagnosis, self-triage, digitalization in primary care, general practitioners; symptom checker; app; mobile app; primary care

Introduction

Background

The number of health-related software in consumer and research-oriented apps is increasing rapidly. Symptom Checker Apps (SCAs) are one example for health-related software that could have a major impact on health systems on all levels. SCAs process medical symptoms that users enter by applying algorithms and databases with medical information [1]. Based on these symptoms, SCAs generate a list of probable causes and suggest medical follow-up actions (eg, wait at home, see a doctor, go to the emergency room). The Google Play Store already lists 249 apps (retrieved on March 24, 2021) for the key words "symptom checker" [2]. Some SCA manufacturers advertise that they have implemented artificial intelligence [1,3,4] and big data as the basis of their apps. For example, as described by Richens et al [5], the SCA "Babylon" uses a causal machine learning approach based on a Bayesian approach combined with counterfactual inference. The presented algorithm achieved expert clinical accuracy for a test set of clinical vignettes [1]. It, however, remains unclear how well it performs in real-life situations. Although there have already been strong claims from ethical research that emphasize the significance of criteria such as transparency, trustworthiness, agency, and responsibility for artificial intelligence–driven decision support systems such as SCAs [6-8], not all manufacturers using artificial intelligence consider these criteria, and it is mainly untransparent how user data are processed and algorithms are trained.

The regulation of SCAs varies in different countries. In some countries, SCAs are effectively unregulated (eg, Australia [9]); in others, SCAs are regulated but with a period of "enforcement discretion" without active regulatory (eg, US Food and Drug Administration [10]). In the European Union, there is a transition between regulation through manufacturer "Declaration" of conformity to legislation toward a model of regulator audit of compliance with standards, including the formal reporting and evaluation of specified forms of clinical data and surveillance trends [11].

Symptom checkers are low-threshold tools that can be accessed with a suitable electronic device with internet access such as a smartphone and are available as apps or as browser versions. Users must be able to interact with technical devices and to interpret the SCA's output to utilize them properly. This could lead to a disadvantage of specific population groups, for example, older adults, people with disabilities [12], or people with limited economic resources [13]. Some SCAs exclude specific user groups for symptom analysis, for example, pregnant women, children, older adults, and patients with specific comorbidities [14].

A recent study by Aboueid et al [15] investigated the intention to use symptom checkers for self-triage and revealed 5 profiles by using a latent class analysis: tech acceptors, tech rejectors, skeptics, unsure acceptors, and tech seekers. Tech seekers, which were described as participants who have positive perspectives related to SCA functionality and artificial intelligence but do not perceive to have access to the technology, showed the highest odds to use SCAs. However, the sample investigated only students aged between 18 and 34 years [15].

Although some users found that SCAs are useful tools for self-diagnosis and even reported positive health effects [16], other users had problems providing and interpreting concrete information on symptom time patterns or severity [17].

SCAs recommend actions and probable causes for the entered symptoms through their output if the output is incongruent with the users' experience or if expectation discrepancies arise [18] and may initiate unnecessary health care–seeking behavior [19].

In terms of their medical value and validity, commercially available SCAs still have problems with accurate triage (determining a user's medical condition based on their input and recommending the optimal health-related actions for the user). Several studies showed that SCAs often suggested risk-averse action recommendations [20-22]. SCA diagnostic and triage accuracy is still limited and was even less reliable in nonurgent scenarios, which are common in primary care [20,23]. A recently published study compared the performance of SCAs (n=8) with the performance of telephone consultation with the general practitioner (GP) (n=7) by using case vignettes (n=100). GPs outperformed SCAs on all assessed outcomes (accuracy, condition suggestion, appropriateness, and safety of urgency advice). The comparison was limited to telephone consultations and did not comprise direct patient-physician contact. Another recent study compared the performance of SCAs to that of medical laypersons using clinical vignettes and found that most laypersons outperformed the majority of SCAs, even though SCAs detected emergency cases more reliably than the laypersons [23].

In high-performing health care systems, inaccurate triage can cause preventable costs and increase the risk of unnecessary procedures that could lead to avoidable risks for patient's safety [3,24]. However, in structurally weak regions with restricted access to medical care, SCAs can provide a first-line assessment that otherwise would not be available [25].

In summary, the potential risks of the use of SCAs (exclusion of users, stress, and induction of health-seeking behavior) contrast the advertised opportunities of SCAs such as patient empowerment and better health care for underserved regions. There is a substantial gap in the literature concerning the effect of SCAs on different health care systems, different levels of

XSL•FO
RenderX

health care (microlevel, mesolevel, and macrolevel [26]) within these systems, and the system's different participants (users, nonusers, and health care providers). If SCAs become more widely used, their ethical, legal, social, and systemic impacts on these levels and participants must be better understood despite complex interactions and methodological challenges. In this study, we aim to clarify the ethical, legal, social, and systemic impacts of SCAs on users, nonusers, GPs, the primary health care systems, and their work by means of an independent, empirical, integrated multi-perspective, and multidisciplinary discussion.

Objectives

Owing to the lack of systematic research of SCAs in primary health care, the recent study uses an explorative hypothesis-generating approach in which the abovementioned discussion is informed by 4 foci of interest and the study aims, as stated in the following section.

Focus 1: Ethical, Legal, and Social Issues of SCA Use

We aim to identify the ethical, social, and legal subjects in the recent scientific literature on SCAs (eg, usage linked to inequities in health care, patient autonomy, modification of role concepts and agency).

Focus 2: SCA Epidemiology, Users, Nonusers, and Predictors of Use

Our results will contribute to describing the user group and nonuser group of SCAs in more detail. The degree of use of SCAs in Germany will be derived, and predictors for SCA use will be identified. Moreover, vulnerable groups that might be disadvantaged through the implementation of SCAs will be described.

Focus 3: Patterns and Impact of SCA Use and the User Level

As a goal for the users' perspective, a comprehensive description of SCA use will be derived. This will comprise the assessment of requirements that are fulfilled or unfulfilled by SCA use. Additionally, we aim to identify the possible risks associated with SCA use and assess how users handle SCA information and action recommendations. SCA effects on user agency, health and eHealth literacy, well-being, and self-care will also be observed.

Focus 4: Impact of SCA Use on Health Care Systems and Health Care Workers

We will specially focus on considering the impact of SCAs from a health system perspective by assessing the following: changes in the patient-physician relationship, strategies of handling preinformed patients, changes in the role concept of physicians and requirements of GP, as well as potential psychosocial risks and demands and perceived work stress resulting from these changes.

Methods

Study Design

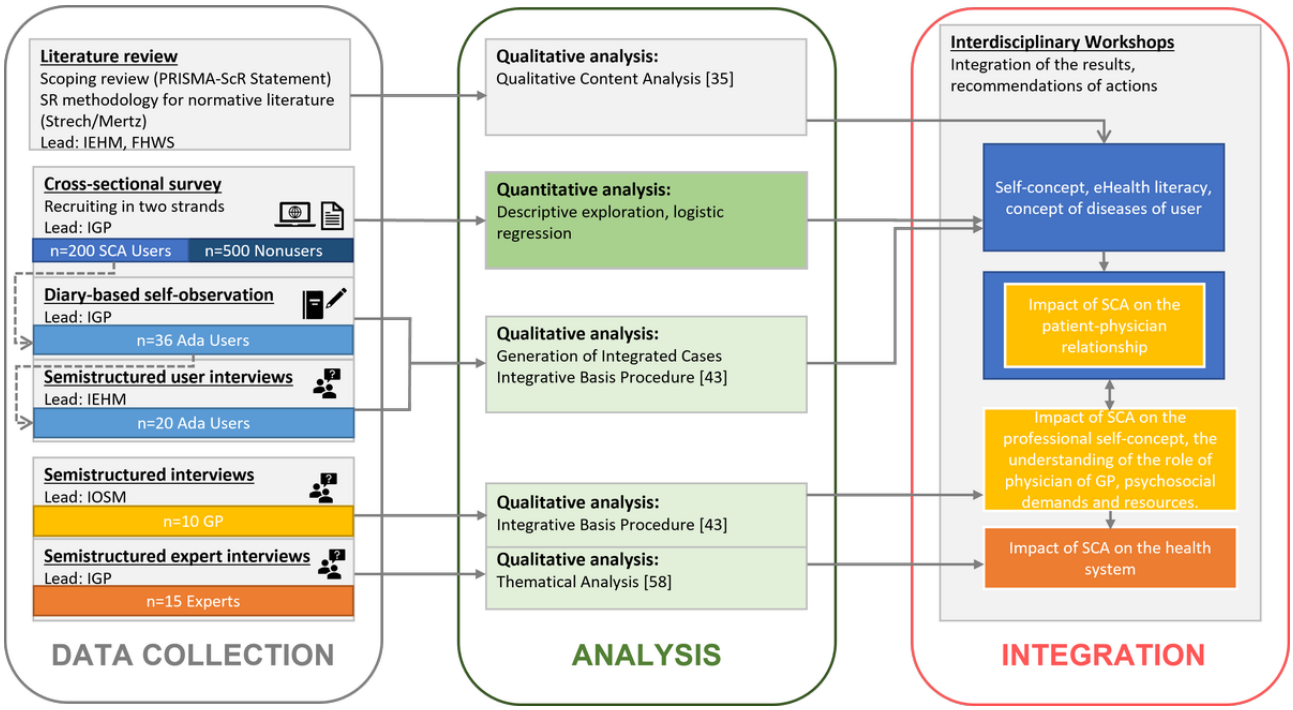
This study's areas of focus and corresponding study phases will be conducted by multidisciplinary partners from the areas of social medicine, ethics and medical history, legal studies, general practice, sociology, occupational health medicine, and health services research. We will investigate SCAs that offer self-triage and action recommendations for medical laypersons. The project partners will co-develop study materials, subsequently discuss results, and conduct method workshops throughout the 3 years. During the final study year, a series of workshops will include participants of the preceding study parts. This workshop series during the final year is led by a social scientist with comprehensive experience of working in cross-disciplinary research and holding method workshops. The workshops will also serve as internal quality control and monitoring. An advisory board is continually informed about the progress and the results of the study phases. Members of the advisory board will be recruited from different contexts and disciplines. The advisory board will meet annually to give feedback on the research process, preliminary results, and the dissemination of the latter.

Ethical Considerations

The German Federal Ministry of Education and Research funds the project for 3 years (Grant 01GP1907A). Ethical approval for this study was obtained from the ethics committee of the University of Tübingen (ID: 464/2020BO). This study will be conducted in accordance with the Declaration of Helsinki. Study participants will be informed thoroughly about the study and their rights, and written informed consent will be obtained from all study participants. Other research ethics requirements such as data protection will be diligently considered. The general study design and the involved research partners are outlined in Figure 1.

XSL•FO
RenderX

Figure 1. Overview of the study design, research partners and analysis methods. IEHM: Institute of Ethics and History of Medicine, University Tübingen, IGP: Institute of General Practice and Interprofessional Care, University Hospital Tübingen, IOSM: Institute of Occupational and Social Medicine and Health Services Research, University Hospital Tübingen, FHWS: University of Applied Sciences Würzburg-Schweinfurt, SCA: Symptom Checker App.



Study Course

Four main data sources will be considered, each representing a specific stakeholder group of SCAs: representative sample of the German population (divided into SCA user and nonuser), GP, and health care experts. We define SCA users as participants that have used SCAs at least once, and nonusers are participants who never used SCAs. Data collection and analysis methods will comprise qualitative and quantitative approaches.

The different methods are applied to the data sources in 4 interconnected study phases, each representing one of the four main foci of interest of the study: a literature review (study phase 1, lead: Institute of Ethics and History of Medicine, University Tübingen [IEHM], Institute of Applied Social Science, University of Applied Science Würzburg-Schweinfurt), a representative survey of SCA user and SCA nonuser (study phase 2, lead: Institute of General Practice and Interprofessional Care, University Hospital Tübingen [IGP]), an SCA user diary-based self-observation combined with individual semistructured interviews (study phase 3, lead: IEHM), and lastly, single semistructured interviews with GPs and health care experts (study phase 4, lead: Institute of Occupational and Social Medicine and Health Services Research, University Hospital Tübingen [IOSM] for GP interviews, lead: IGP for health care experts interviews). For this study protocol, we will follow the Good Reporting of A Mixed Methods Study (GRAMMS) [27] checklist. Preferred Reporting Items for Systematic reviews and Meta-Analyses extension for Scoping Reviews (PRISMA-ScR) [28,29], STROBE (STrengthening the Reporting of OBservational studies in Epidemiology) [30], SRQR (Standards for Reporting Qualitative Research) [31], and GRAMMS [28] checklists will be applied during the project process. The PRISMA-ScR, STROBE, and SRQR guidelines will be applied on specific study phases. The GRAMMS [28] guideline will be used in the context of mixed methods approaches and in the integration of results. In the following sections, the 4 study phases and their connections are described in detail.

Study Phase 1: Literature Review

The first study phase will comprise a comprehensive literature review that will assess the existing research on SCAs and their impact on primary care. The aim of this study phase is to gain a clearer picture of the state of science of SCAs, considering the ethical, legal, social, and systemic (eg, risks, potentials) impacts of SCAs. A literature search, oriented on scoping review (ScR) methodology, will be conducted and reported according to the PRISMA-ScR statement [28,29]. In recent years, the ScR methodology has been adopted and further developed for the field of bioethics, which is characterized by normative research questions. When analyzing argumentative literature, adjustments need to be made to the “classic” ScR methodology [32-34]. Key terms will be defined for the search strategy regarding the research questions, and databases covering the relevant dimensions (biomedical, ethical, social, and legal) are selected (Web of Science, PubMed [Medline], Belit/Ethmed, ProQuest, SowiPort, GESIS, Philpapers, Juris, BeckOnline, etc). The inclusion criteria comprise either the mention of an ethical, legal, or social issue or a combination of this terms related to SCAs as a digital or mobile app, targeting medical laypersons that support the assessment of symptoms and self-triage or one of both aspects. Apps for health care professionals or other experts, as well as SCAs, which focus on a single health condition or certain groups of diseases were excluded. To present the full spectrum of ethical, legal, social, and systemic impacts literature relevant to the review question, not only

XSL•FO
RenderX

argument-based but also empirical literature was included. Publications on SCAs that were written in English or German and published until 2020 were included in the review. Journal articles, contributions to anthologies, reports, case series, letters to editors, opinions, commentaries, and conference papers were also included; web-based blogs have not been considered. Three researchers screened the identified literature via hand and database search and discarded publications not meeting the inclusion criteria. Publications were analyzed by 2 authors using the method of qualitative content analysis proposed by Kuckartz [35].

Study Phase 2: Cross-sectional Survey of SCA Users and SCA Nonusers

Structure

A survey will be used in a case-control design. The questionnaire was piloted with 5 participants and will take approximately 15-25 minutes for participants to complete. The survey will comprise different evaluated scales (for further details, see Multimedia Appendix 1) and sociodemographic variables as well as specific questions to the usage of SCAs. This study phase will be conducted and reported according to the STROBE statement [30]. Owing to the limited amount of literature on SCA users and nonusers, pilot interviews with 2 SCA users and 1 SCA expert will be conducted to ensure a meaningful concept selection for the survey. Simultaneously, concepts will be derived from existing literature that is connected to the use of health apps and could reveal the potential characteristics of the user group such as eHealth Literacy [36], personality [37,38], hypochondria [39], self-efficacy [40,41], and need for cognition [42]. Affinity for technology [43], satisfaction with the GP [44], and overall life satisfaction [45] will also be considered. Moreover, we will assess the perceived usefulness of SCAs and requirements for SCAs in open-ended questions of the survey since this might be a central aspect of acceptance of SCAs as the Technology Acceptance Model (TAM) induces. The TAM was already introduced by Davis [46] in 1989 and is based on the theory of reasonable action and the theory of planned behavior. The TAM (and its further expansions) is still one of the most prevailing models to examine factors affecting a users' acceptance of new technologies [47]. The model assumes a mediating function of perceived ease of use and usefulness in association with system characteristics and system usage [47].

Sample and Recruitment

The sampling process was conducted from November 2020 until the end of June 2021. A case-control design using 2 recruitment strands is planned. An a priori power analysis using PASS 2020 (v20.0.3, NCSS) revealed a sample size of 375 (β=.8, α=.05, nuisance parameter=0.2; n_{user}=188, $n_{nonuser}$=188) for an odds ratio of 2.5. The targeted odds ratio corresponds to a small-to-medium effect of Cohen *d* [48] and was selected because we consider that this will be an effect size that contributes a meaningful explanation of variance in the logistic regression. As this study has an explorative character, we could not derive theory-driven assumptions for a multivariate logistic regression. Hence, we based the power analysis on a univariate logistic regression. We will use a univariate logistic regression to identify meaningful predictors for the usage of SCAs and will moreover set up a multivariate model that includes all the identified predictors. Our multivariate model will be the first proposal and will need further research to confirm the univariate predictors in a multivariate model.

The sample will be composed of different recruiting strands to achieve a representative sample. In the first strand, German citizens will be contacted via mail to participate in the survey. The intended recipients will be representatively selected by an external partner (T + R Dialog Marketing and Acxiom). Further participants will be recruited via mailing lists of the University of Tübingen and the University Hospital of Tübingen, social media, and cooperating GP practices. After 3 months, the representativity of this sample will be checked. If the return rate is too low or if certain groups are not sufficiently represented, there will be additional recruiting via the proposed channels.

The second strand of the sampling process aims to integrate symptom checker users only. We expect only a small number of symptom checker users. To ensure a sample size of n_{user}=188, a targeted recruitment via social media advertisements and the social media channels of the University Hospital of Tübingen will be conducted.

Inclusion criteria in general are the ability to give consent and German language skills of at least B1 of the Common European Framework of References for Languages. Participants of the second recruitment strand can only be included if they have experience with SCAs.

Analysis

The level of use for SCAs, awareness of SCAs, and general interest in SCAs will be described using the first recruitment strand with descriptive statistics.

Following the case-control design, SCA users from the first and second recruiting strand will be matched with nonusers (matched controls) from the first recruitment strand. Significant predictors will be extracted with a logistic regression. A correction for multiple testing will be applied. The recent versions of SPSS (IBM) and R statistic (R Core Team) will be used for the analyses.

Study Phase 3: Diary-Based Self-observation Combined With Semistructured User Interviews

Structure

Study phase 3 investigates SCA users and their usage patterns and effects of SCAs on individuals. A specific SCA (Ada app) was chosen by the study team as an example since it is considered one of the most prevalent SCAs in Germany. Following web-based training, participants will engage in a diary-based self-observation. During the observation time of 6 weeks, participants will document their daily usage and nonusage of the Ada app. Next, individual semistructured user interviews are performed with the diary study participants. The interviews allow participants to reflect on values, concepts, and knowledge gaps. This allows a supplemental exploration of the experiences recorded in the diaries. The user interviews in this

XSL•FO
RenderX

study part will be conducted and reported according to SRQR statement [31].

Sample and Recruitment

For this study phase, 50 Ada users will be recruited from the SCA user strand of phase 2 and, if needed, additionally via social media. Considering a dropout rate of 30%, a sample size of 36 is assumed. Of these participants, 15 will be recruited for single semistructured interviews using maximum variation sampling. Sampling will consider the content of the diary-based self-observation, usage behavior of the app, medical indication, and socioeconomic factors. Sample size calculations of the interview phases are based on the 5D model of *information power* by Malterud et al [49].

All participants will receive web-based training on the self-observation period. The diary will be used to document symptoms and events, as well as other expected influencing factors such as stress or quality of life. Furthermore, it will offer structured questions about the use of SCAs and enable the participants to write down their own reports or short "field notes." Thus, participants will record and describe their experience, and how they dealt with action recommendations, appearing problems, emotions, etc. These notes will be used as a basis for the following semistructured interviews. The interviews will be conducted via video call, audio-recorded, and transcribed verbatim by a researcher from the IGP and the IEHM. The users will receive financial compensation both for the interviews and the participation in the member check meeting described below.

Analysis

The diary-based self-observation and the interview transcripts will be analyzed and integrated into cases. Triangulating the self-observation diary data with interview transcripts provides both prospective (longitudinal) and retrospective (narrative) insights. Quantitative results of the diaries (frequency of use, use of health care, symptoms, etc) will be considered as prospective observational outcomes and will be analyzed quantitatively. The user diaries give the opportunity to record detailed situational experiences—feelings and thoughts that probably cannot be remembered or recreated during an interview without them. However, the interviews will provide in-depth reports on values, concepts, gaps in knowledge, etc, that tend to remain invisible in the diaries. The qualitative analysis via Kruse's integrative basic procedure [43] will provide an overview of the recurring themes and patterns within each case as well as between cases. At the same time, it will allow a more holistic consideration of the data such as the analysis of semantics, grammatical structures, and metaphors to reveal latent meanings and the way users "make sense" of the app and derive meaning and understanding of the recorded events [50]. The qualitative analysis of the study is supported by MaxQDA [51]. A member check with participants of the interviews is planned, in which results of the cases are presented to study participants to enhance rigor.

The aim of the quantitative analysis is to identify meaningful predictors for the use of SCAs, taking the longitudinal data structure into account. A hierarchical model with 2 levels will be performed. Level two will comprise the daily measurements and will be nested in level one, which comprises the participants. The quantitative analysis of the diaries will be performed using a recent version of Microsoft Excel [52] and R Statistics (R Core Team).

Study Phase 4: Semistructured Interviews With GPs and Health Care Experts

Structure

The fourth study phase investigates the possible effects of SCAs on health care delivery, health care providers (module A), and the health care system (module B). As primary care is most affected by patients' usage of SCAs, we will interview GPs in module A. We will gain more insights into patients' usage of SCAs and similar application results in potential psychosocial demands, resources, and perceived work-related stress [53], especially regarding workload, work content, work organization, and social environment [54]. Module B aims to deliver a multi-perspective view on possible effects of SCAs on the health care system. To fulfill this aim, we will conduct interviews with health care experts with different backgrounds to assess the state of science of SCAs in practice from a multi-perspective standpoint. Moreover, we aim to identify which potential experts see for the future use of SCAs and to derive quality criteria for SCAs. This study phase will be conducted and reported according to SRQR statement [31].

Sample and Recruitment

In module A, the sample will consist of 10 GPs in Germany. We aim to build a heterogeneous sample regarding the GP (age, gender, and race) and their practices (structure and location of the practice, main patient clientele, and availability of web-based services). Sample size calculations of the interview phases in modules A and B are based on the 5D model of *information power* by Malterud et al [49]. An interview guide [55] will be developed, containing questions about preinformed patients and diagnosis in general, questions about SCAs, and similar apps with the example of the Ada app in particular. It will be developed by the IOSM and the help of feedback from the other project partners. The IOSM will apply various forms of sampling such as snowball sampling in the established networks of the IOSM and web-based research to ensure the stated heterogeneity of the sample. Two researchers of the IOSM will conduct the interviews mostly via video call due to the ongoing pandemic. The interviews are expected to last about 45 minutes and will be audio-recorded.

In module B, 10-15 experts on health care systems will be interviewed. Experts will be recruited consecutively and comprise politicians, information technology developers for medical software, patient advocates, representatives from jurisdiction, medical associations, and health insurances. We assume that 10-15 interviews will provide sufficient information power [49]. Each interview guide will be tailored to the respective expert. Possible topics are the implementation of SCAs, recent issues with SCA requirements, and how SCAs influence different players in health care. The interview guides will be developed by the team of the IGP with input from the other project partners. The experts will be contacted via already

XSL•FO
RenderX

existing research networks of the project partners. One researcher of the IGP will conduct the interviews mostly via video call. The interviews are planned to last about 45-90 minutes and will be audio recorded.

The expert interviews will provide an information background on the *status quo* of SCAs in health care as well as ideas for future developments. This background is important for the discussion of user experiences (Study phase 3) and for the patient-physician relationship (integrated study phases 3 and 4). All participants will receive financial compensation both for the interviews and the participation in the member check meeting.

Analysis

All interviews will be transcribed. In module A, the interviews will be analyzed with reflexive thematical analysis by Braun and Clark [56]. As already stated, this method is used to understand how interview participants "make sense" of their experiences. This allows insight into tacit changes in self-concepts, implicit values psychosocial demands and resources, and perceived work-related stress. Additionally, it allows to conduct an analysis of themes and content provided by the interview partners. The IOSM team will individually evaluate each interview and compare the interviews to analyze common patterns. To ensure quality control and richness of analysis, each interview will be analyzed by 2 researchers, and preliminary results will be continuously discussed with additional researchers from the joint project. The interviews in module B will be analyzed with reflexive thematical analysis by Braun and Clark [56], as we aim to collect and structure the overarching themes and their various dimensions. The same measures for quality control, as described in module A, will be applied. For additional quality control, a member check with interview participants of phase 4 is planned: preliminary results will be presented, and participants will be invited to give critical feedback, which will be integrated into the further analysis.

Results

The project started in March 2020.

Study Phase 1: Literature Review

The literature review on ethical, legal, social, and systemic impacts of SCAs was completed in December 2021 and will be published in 2022.

Study Phase 2: Cross-sectional Survey of SCA Users and SCA Nonusers

The data collection for study phase 2 was finished in July 2021 (n=1074); the publication of the results is planned for 2022.

Study Phase 3: Diary-Based Self-observation Combined With Semistructured User Interviews

For study phase 3, data collection of the user diaries (n=48) was completed in October 2021, user interviews were conducted in February and March 2022, and publication of the results is planned for 2022.

Study Phase 4: Semistructured Interviews With GPs and Health Care Experts

The recruitment of experts and GPs (study phase 4) is still ongoing and will be finished in 2022.

Discussion

Principal Findings

This protocol describes an interdisciplinary, mixed methods, multiphase research program to comprise the impact of SCAs on the 3 levels (microlevel, mesolevel, and macrolevel) of the health care system [26]. The main findings of the recent exploratory study will be an overview of the ethical, legal, social, and systemic impacts literature on SCAs, epidemiology data of SCAs (as degree of use in Germany, predictors for SCA usage considering user characteristics), requirements of SCA use from a user perspective, a description of SCA user behavior, and a comprehensive assessment of the perspective of GPs and health care experts on SCAs.

Digital health care innovations and their impact on the health care system is a prevailing topic in times of digitalization and the increasing demand of health care professionals such as GPs. In contrast to apps for specialists, user-accessible apps such as SCAs are an unknown variable in the development of health care delivery in the future. Conflicting claims of medical, ethical, and social advantages or disadvantages of SCAs characterize the current state of the debate. Apparently paradoxical effects, as undermining trust in the patient-physician relationship on the one hand, improved exchange on the other hand, may be coexisting or representing different perspectives in scenarios that require further description. The conflicting information about ethical, legal, social, and systemic impacts requires more empiric data to inform and deepen the debate [57].

Little is known about the psychosocial demands and resources of GPs in this context. SCAs are attributed to result in an overuse as well as in an underuse of health care resources. Finally, existing health inequalities may be improved or worsened by their impact. Based on an ScR of the ethical, social, and legal literature, the project will provide evidence, which of these contradicting assumptions is confirmed by the empirical study of the user experiences in the case study at hand. Using SCA as an example of consumer-oriented digital innovations, this study aims to research and integrate questions that are important for the general debate on digital transformation: what is known about the topic? How widespread is the phenomenon? How do users apply the innovation? How does that modulate their behavior and impact their health care usage? How do negotiations with health care providers play out? Which regulatory legislation is necessary? What are the implications for the physicians in particular and the health care system as a whole? However, only collecting data from 1 stakeholder's perspective without considering possible interactions will generate blind spots. Thus, the main challenge is to consider different stakeholders' perspectives, wants and needs, and to engage in a transparent debate on the current dynamic developments.

XSL•FO
RenderX

Strengths and Limitations

The projects' concept integrates different data sources and methods from the very start. The multidisciplinary, multiphase design, and the methods and skills mix of the study partners create a scenario in which methodological strengths are complimentary and perspectives can be negotiated. For example, by limiting recruitment to a single SCA (the Ada app) in the qualitative study phases, we are able to focus the analysis [49]. At the same time, through the representative survey, these qualitative results can be put in a broader context, which will contribute to implications for the health care system. Another strength is the immanent consideration of the user perspective by combining a survey, user diary, and user interviews. In integrating relevant perspectives and plotting study phases to converge in integrated workshops, we present an approach for integrated research in ethics, social, and health sciences. This is possible owing to the long-standing cooperation between all involved partners and previous positive experiences in common projects.

This research project also has limitations: we will not investigate how SCAs perform in terms of medical accuracy. Further, we will not be able to observe direct interaction of patients and physicians in the context of SCAs and see how SCAs influence the patient-physician relationship directly. Moreover, we assessed user and nonuser characteristics through subjective user rating rather than objective measurements (eg, overall health rating of participant). Lastly, we will not directly investigate the effect of SCAs on health care utilization. We will, however, assess parameters considering utilization reported by participants.

Conclusions

This study offers an opportunity for multidisciplinary research: it considers different research perspectives and methodologies from ethics, legal, social, health care, and medical science and integrates them in 1 study process. We are confident that this will lead to new insights for the use of SCAs and digitalization in health care while providing a novel methodological approach for integrated research in health care digitalization.

Acknowledgments

We thank Anne-Kristin Münch and Prof Dr Peter Martus for their advice. Further, we thank Marius Wetzel for the language check of the manuscript. The project proposal of the research project CHECK.APP has undergone an external peer review process before it was accepted for funding. The project is fully funded by the German Federal Ministry of Education and Research who have no role in the design of the study phases and will have no ruling during its execution, or in the analyses of the data, or in the decision to submit any results. This study protocol has not undergone peer review by the funding body. Contact information: Heinrich-Kronen-Straße 1, 43227 Bonn; contact person for CHECK.APP: Dr. Katja Hüttner; email: Katja.Huettner@dlr.de.

Authors' Contributions

AJW and RK equally contributed to the manuscript draft and the writing process. RK, SJ, MAR, UW, HJE, CP, RR, and TH collaborated on the research proposal grant, study protocol, and study design. The authors' contributions to the design of the study phases and the editing of the corresponding phases in the manuscript were as follows: UW, RM, RR, HJE, TH, and MK contributed to study phase 1. AJW, SJ, and RK contributed to study phase 2. Phase 3 was collaboratively contributed to by RK, CP, MK, AJW, and RM. MAR and CP contributed to phase 4A; AJW, RK, and CP contributed to phase 4B. All authors collaborated on the editing of the manuscript draft. All authors approved the final version.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Evaluated scales that were used in the survey of study phase 2.
[PNG File , 77 KB-Multimedia Appendix 1]

References

1. Richens JG, Lee CM, Johri S. Improving the accuracy of medical diagnosis with causal machine learning. Nat Commun 2020 Aug 11;11(1):3923 [FREE Full text] [doi: 10.1038/s41467-020-17419-7] [Medline: 32782264]
2. Google Play Store. URL: https://play.google.com/store?gl=DE [accessed 2022-05-06]
3. Aboueid S, Liu RH, Desta BN, Chaurasia A, Ebrahim S. The Use of Artificially Intelligent Self-Diagnosing Digital Platforms by the General Public: Scoping Review. JMIR Med Inform 2019 May 01;7(2):e13445 [FREE Full text] [doi: 10.2196/13445] [Medline: 31042151]
4. TK legt Kooperation mit Ada Health auf Eis. Aerztezeitung. 2019. URL: https://www.aerztezeitung.de/Wirtschaft/TK-legt-Kooperation-mit-Ada-Health-auf-Eis-404721.html [accessed 2020-05-15]
5. Richens JG, Lee CM, Johri S. Publisher Correction: Improving the accuracy of medical diagnosis with causal machine learning. Nat Commun 2020 Sep 16;11(1):4754 [FREE Full text] [doi: 10.1038/s41467-020-18310-1] [Medline: 32938913]
6. Braun M, Hummel P, Beck S, Dabrock P. Primer on an ethics of AI-based decision support systems in the clinic. J Med Ethics 2020 Apr 03 [FREE Full text] [doi: 10.1136/medethics-2019-105860] [Medline: 32245804]

XSL•FO
RenderX

7. Jobin A, Ienca M, Vayena E. The global landscape of AI ethics guidelines. Nat Mach Intell 2019 Sep 02;1(9):389-399. [doi: 10.1038/s42256-019-0088-2]
8. Ienca M, Ferretti A, Hurst S, Puhan M, Lovis C, Vayena E. Considerations for ethics review of big data health research: A scoping review. PLoS One 2018;13(10):e0204937 [FREE Full text] [doi: 10.1371/journal.pone.0204937] [Medline: 30308031]
9. Australian Government - Department of Health. Regulation of software based medical devices. URL: https://www.tga.gov.au/regulation-software-based-medical-devices [accessed 2021-02-01]
10. U.S. Department of Health and Human Services. Policy for Device Software Functions and Mobile Medical Applications. 2019. URL: https://www.fda.gov/media/80958/download [accessed 2022-02-18]
11. European parliament and the council of the European Union. REGULATION (EU) 2017/745 OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL. Official Journal of the European Union. 2017. URL: https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32017R0745 [accessed 2022-02-18]
12. Hageman MGJS, Anderson J, Blok R, Bossen JKJ, Ring D. Internet self-diagnosis in hand surgery. Hand (NY) 2015 Sep;10(3):565-569 [FREE Full text] [doi: 10.1007/s11552-014-9707-x] [Medline: 26330798]
13. Scheuer S. Warum die Corona-App auf Millionen Smartphones nicht funktioniert. Handelsblatt. URL: https://www.handelsblatt.com/technik/it-internet/alte-betriebssysteme-warum-die-corona-app-auf-millionen-smartphones-nicht-funktioniert/25929842.html?ticket=ST-2591858-cKGesa05dbuV5H5br12G-ap5 [accessed 2021-04-28]
14. Gilbert S, Mehl A, Baluch A, Cawley C, Challiner J, Fraser H, et al. How accurate are digital symptom assessment apps for suggesting conditions and urgency advice? A clinical vignettes comparison to GPs. BMJ Open 2020 Dec 16;10(12):e040269 [FREE Full text] [doi: 10.1136/bmjopen-2020-040269] [Medline: 33328258]
15. Aboueid S, Meyer SB, Wallace J, Chaurasia A. Latent classes associated with the intention to use a symptom checker for self-triage. PLoS One 2021;16(11):e0259547 [FREE Full text] [doi: 10.1371/journal.pone.0259547] [Medline: 34731217]
16. Meyer AND, Giardina TD, Spitzmueller C, Shahid U, Scott TMT, Singh H. Patient Perspectives on the Usefulness of an Artificial Intelligence-Assisted Symptom Checker: Cross-Sectional Survey Study. J Med Internet Res 2020 Jan 30;22(1):e14679 [FREE Full text] [doi: 10.2196/14679] [Medline: 32012052]
17. Marco-Ruiz L, Bønes E, de la Asunción E, Gabarron E, Aviles-Solis JC, Lee E, et al. Combining multivariate statistics and the think-aloud protocol to assess Human-Computer Interaction barriers in symptom checkers. J Biomed Inform 2017 Oct;74:104-122 [FREE Full text] [doi: 10.1016/j.jbi.2017.09.002] [Medline: 28893671]
18. Luger TM, Houston TK, Suls J. Older adult experience of online diagnosis: results from a scenario-based think-aloud protocol. J Med Internet Res 2014 Jan 16;16(1):e16 [FREE Full text] [doi: 10.2196/jmir.2924] [Medline: 24434479]
19. Winn AN, Somai M, Fergestrom N, Crotty BH. Association of Use of Online Symptom Checkers With Patients' Plans for Seeking Care. JAMA Netw Open 2019 Dec 02;2(12):e1918561 [FREE Full text] [doi: 10.1001/jamanetworkopen.2019.18561] [Medline: 31880791]
20. Semigran HL, Linder JA, Gidengil C, Mehrotra A. Evaluation of symptom checkers for self diagnosis and triage: audit study. BMJ 2015;351:h3480-h3489 [FREE Full text] [Medline: 26157077]
21. Hill MG, Sim M, Mills B. The quality of diagnosis and triage advice provided by free online symptom checkers and apps in Australia. Med J Aust 2020 Jun;212(11):514-519. [doi: 10.5694/mja2.50600] [Medline: 32391611]
22. Ceney A, Tolond S, Glowinski A, Marks B, Swift S, Palser T. Accuracy of online symptom checkers and the potential impact on service utilisation. PLoS One 2021;16(7):e0254088 [FREE Full text] [doi: 10.1371/journal.pone.0254088] [Medline: 34265845]
23. Schmieding ML, Mörgeli R, Schmieding MAL, Feufel MA, Balzer F. Benchmarking Triage Capability of Symptom Checkers Against That of Medical Laypersons: Survey Study. J Med Internet Res 2021 Mar 10;23(3):e24475 [FREE Full text] [doi: 10.2196/24475] [Medline: 33688845]
24. Marckmann G. Ethische Aspekte von eHealth. In: Fischer F, Krämer A, editors. eHealth in Deutschland. Berlin: Springer Verlag; 2016:83-99.
25. Fiske A, Buyx A, Prainsack B. The double-edged sword of digital self-care: Physician perspectives from Northern Germany. Soc Sci Med 2020 Sep;260:113174 [FREE Full text] [doi: 10.1016/j.socscimed.2020.113174] [Medline: 32659512]
26. Leijten FRM, Struckmann V, van Ginneken E, Czypionka T, Kraus M, Reiss M, SELFIE consortium. The SELFIE framework for integrated care for multi-morbidity: Development and description. Health Policy 2018 Jan;122(1):12-22 [FREE Full text] [doi: 10.1016/j.healthpol.2017.06.002] [Medline: 28668222]
27. O'Cathain A, Murphy E, Nicholl J. The quality of mixed methods studies in health services research. J Health Serv Res Policy 2008 Apr;13(2):92-98. [doi: 10.1258/jhsrp.2007.007074] [Medline: 18416914]
28. Liberati A, Altman DG, Tetzlaff J, Mulrow C, Gøtzsche PC, Ioannidis JPA, et al. The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate healthcare interventions: explanation and elaboration. BMJ 2009 Jul 21;339:b2700 [FREE Full text] [doi: 10.1136/bmj.b2700] [Medline: 19622552]
29. Tricco AC, Lillie E, Zarin W, O'Brien KK, Colquhoun H, Levac D, et al. PRISMA Extension for Scoping Reviews (PRISMA-ScR): Checklist and Explanation. Ann Intern Med 2018 Sep 04;169(7):467. [doi: 10.7326/M18-0850]

XSL•FO
RenderX

30. von Elm E, Altman DG, Egger M, Pocock SJ, Gøtzsche PC, Vandenbroucke JP, STROBE Initiative. The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) Statement: guidelines for reporting observational studies. Int J Surg 2014 Dec;12(12):1495-1499 [FREE Full text] [doi: 10.1016/j.ijsu.2014.07.013] [Medline: 25046131]
31. O'Brien BC, Harris IB, Beckman TJ, Reed DA, Cook DA. Standards for reporting qualitative research: a synthesis of recommendations. Acad Med 2014 Sep;89(9):1245-1251 [FREE Full text] [doi: 10.1097/ACM.0000000000000388] [Medline: 24979285]
32. Strech D, Synofzik M, Marckmann G. Systematic reviews of empirical bioethics. J Med Ethics 2008 Jun;34(6):472-477. [doi: 10.1136/jme.2007.021709] [Medline: 18511623]
33. Strech D, Sofaer N. How to write a systematic review of reasons. J Med Ethics 2012 Feb;38(2):121-126 [FREE Full text] [doi: 10.1136/medethics-2011-100096] [Medline: 22080465]
34. Mertz M, Kahrass H, Strech D. Current state of ethics literature synthesis: a systematic review of reviews. BMC Med 2016 Oct 03;14(1):152 [FREE Full text] [doi: 10.1186/s12916-016-0688-1] [Medline: 27716264]
35. Kuckartz U. Qualitative text analysis. Thousand Oaks, California: SAGE Publications Ltd; 2013.
36. Bol N, Helberger N, Weert JCM. Differences in mobile health app use: A source of new digital inequalities? The Information Society 2018 Apr 26;34(3):183-193. [doi: 10.1080/01972243.2018.1438550]
37. Carvalho LF, Sette CP, Ferrari BL. Problematic smartphone use relationship with pathological personality traits: Systematic review and meta-analysis. Cyberpsychology: Journal of Psychosocial Research on Cyberspace 2018 Dec 31;12(3):1-20. [doi: 10.5817/cp2018-3-5]
38. Svendsen GB, Johnsen JK, Almås-Sørensen L, Vittersø J. Personality and technology acceptance: the influence of personality factors on the core constructs of the Technology Acceptance Model. Behaviour & Information Technology 2013 Apr;32(4):323-334. [doi: 10.1080/0144929X.2011.553740]
39. Loos A. Cyberchondria: Too Much Information for the Health Anxious Patient? Journal of Consumer Health On the Internet 2013 Oct;17(4):439-445. [doi: 10.1080/15398285.2013.833452]
40. Balapour A, Reychav I, Sabherwal R, Azuri J. Mobile technology identity and self-efficacy: Implications for the adoption of clinically supported mobile health apps. International Journal of Information Management 2019 Dec;49:58-68. [doi: 10.1016/j.ijinfomgt.2019.03.005]
41. Zhang X, Han X, Dang Y, Meng F, Guo X, Lin J. User acceptance of mobile health services from users' perspectives: The role of self-efficacy and response-efficacy in technology acceptance. Inform Health Soc Care 2017 Mar;42(2):194-206. [doi: 10.1080/17538157.2016.1200053] [Medline: 27564428]
42. Nikoloudakis IA, Crutzen R, Rebar AL, Vandelanotte C, Quester P, Dry M, et al. Can you elaborate on that? Addressing participants' need for cognition in computer-tailored health behavior interventions. Health Psychol Rev 2018 Dec;12(4):437-452. [doi: 10.1080/17437199.2018.1525571] [Medline: 30221580]
43. Kruse J, Schmieder C. Qualitative Interviewforschung: ein integrativer Ansatz. Weinheim: Beltz Juventa; 2014.
44. Scholl I, Hölzel L, Härter M, Dierks M, Bitzer E, Kriston L. Fragebogen zur Zufriedenheit in der ambulanten Versorgung – Schwerpunkt Patientenbeteiligung (ZAPA). Klinische Diagnostik und Evaluation 2011:50-62. [Medline: 1864-6050]
45. Beierlein C, Kovaleva A, László Z, Kemper C, Rammstedt B. Kurzskala zur Erfassung der Allgemeinen Lebenszufriedenheit (L-1). Zusammenstellung sozialwissenschaftlicher Items und Skalen (ZIS) 2015. [doi: 10.6102/zis229]
46. Davis FD. Perceived Usefulness, Perceived Ease of Use, and User Acceptance of Information Technology. MIS Quarterly 1989 Sep;13(3):319. [doi: 10.2307/249008]
47. Rahimi B, Nadri H, Lotfnezhad Afshar H, Timpka T. A Systematic Review of the Technology Acceptance Model in Health Informatics. Appl Clin Inform 2018 Jul;9(3):604-634 [FREE Full text] [doi: 10.1055/s-0038-1668091] [Medline: 30112741]
48. Chen H, Cohen P, Chen S. How Big is a Big Odds Ratio? Interpreting the Magnitudes of Odds Ratios in Epidemiological Studies. Communications in Statistics - Simulation and Computation 2010 Mar 31;39(4):860-864. [doi: 10.1080/03610911003650383]
49. Malterud K, Siersma VD, Guassora AD. Sample Size in Qualitative Interview Studies: Guided by Information Power. Qual Health Res 2016 Nov;26(13):1753-1760. [doi: 10.1177/1049732315617444] [Medline: 26613970]
50. Bruner J. Acts of Meaning. Cambridge, Massachusetts: Harvard University Press; 1990.
51. MAXQDA 2022: The Game Changer. MAXQDA. URL: https://www.maxqda.de/ [accessed 2022-05-06]
52. Microsoft. URL: https://www.microsoft.com/ [accessed 2022-05-06]
53. Demerouti E, Bakker AB, Nachreiner F, Schaufeli WB. The job demands-resources model of burnout. J Appl Psychol 2001 Jun;86(3):499-512. [Medline: 11419809]
54. Beck D, Lenhardt U. Consideration of psychosocial factors in workplace risk assessments: findings from a company survey in Germany. Int Arch Occup Environ Health 2019 Apr;92(3):435-451 [FREE Full text] [doi: 10.1007/s00420-019-01416-5] [Medline: 30756179]
55. Helfferich C. Die Qualität qualitativer Daten. In: Wiesbaden ISBN. Wiesbaden: Springer VS Verlag für Sozialwissenschaften; 2011.
56. Braun V, Clarke V. What can "thematic analysis" offer health and wellbeing researchers? Int J Qual Stud Health Well-being 2014;9:26152 [FREE Full text] [doi: 10.3402/qhw.v9.26152] [Medline: 25326092]

XSL•FO
RenderX

57. Lupton D, Jutel A. 'It's like having a physician in your pocket!' A critical analysis of self-diagnosis smartphone apps. Soc Sci Med 2015 May;133:128-135. [doi: 10.1016/j.socscimed.2015.04.004] [Medline: 25864149]

Abbreviations

GP: general practitioner
GRAMMS: Good Reporting of A Mixed Methods Study
IEHM: Institute of Ethics and History of Medicine, University Tübingen
IGP: Institute of General Practice and Interprofessional Care, University Hospital Tübingen
IOSM: Institute of Occupational and Social Medicine and Health Services Research, University Hospital Tübingen
PRISMA-ScR: Preferred Reporting Items for Systematic reviews and Meta-Analyses extension for Scoping Reviews
SCA: Symptom Checker App
ScR: scoping review
SRQR: Standards for Reporting Qualitative Research
STROBE: STrengthening the Reporting of OBservational studies in Epidemiology
TAM: Technology Acceptance Model

Edited by T Leung; submitted 05.10.21; peer-reviewed by R Cochran, S Aboueid, S Gilbert; comments to author 06.02.22; revised version received 18.02.22; accepted 30.03.22; published 16.05.22

Please cite as:
Wetzel AJ, Koch R, Preiser C, Müller R, Klemmt M, Ranisch R, Ehni HJ, Wiesing U, Rieger MA, Henking T, Joos S
Ethical, Legal, and Social Implications of Symptom Checker Apps in Primary Health Care (CHECK.APP): Protocol for an Interdisciplinary Mixed Methods Study
JMIR Res Protoc 2022;11(5):e34026
URL: https://www.researchprotocols.org/2022/5/e34026
doi: 10.2196/34026
PMID:

©Anna-Jasmin Wetzel, Roland Koch, Christine Preiser, Regina Müller, Malte Klemmt, Robert Ranisch, Hans-Jörg Ehni, Urban Wiesing, Monika A Rieger, Tanja Henking, Stefanie Joos. Originally published in JMIR Research Protocols (https://www.researchprotocols.org), 16.05.2022. 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 work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on https://www.researchprotocols.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

2.2. Publication 2

Publication 2, titled "Better see a doctor? Status Quo of Symptom Checker Apps in Germany: A Cross-Sectional Survey with Mixed-Method Design (CHECK.APP)," examines the prevalence of SCA usage among laypeople in Germany, explores reasons for intentional non-use, and investigates associations between SCA use and primary care visits. This research was published in the journal Sage Digital Medicine (Impact Factor 2023: 3.9).

Check for updates

Original Research Article

DIGITAL HEALTH

DIGITAL HEALTH
Volume 10: 1–11
© The Author(s) 2024
Article reuse guidelines:
sagepub.com/journals-permissions
DOI: 10.1177/20552076241231555
journals.sagepub.com/home/dhj

Sage

‘Better see a doctor?’ Status quo of symptom checker apps in Germany: A cross-sectional survey with a mixed-methods design (CHECK.APP)

Anna-Jasmin Wetzel[1], Roland Koch[1], Nadine Koch[2], Malte Klemmt[3], Regina Müller[4], Christine Preiser[5], Monika Rieger[5], Inka Rösel[6], Robert Ranisch[7], Hans-Jörg Ehni[8] and Stefanie Joos[1]

Abstract

Background: Symptom checker apps (SCAs) offer symptom classification and low-threshold self-triage for laypeople. They are already in use despite their poor accuracy and concerns that they may negatively affect primary care. This study assesses the extent to which SCAs are used by medical laypeople in Germany and which software is most popular. We examined associations between satisfaction with the general practitioner (GP) and SCA use as well as the number of GP visits and SCA use. Furthermore, we assessed the reasons for intentional non-use.

Methods: We conducted a survey comprising standardised and open-ended questions. Quantitative data were weighted, and open-ended responses were examined using thematic analysis.

Results: This study included 850 participants. The SCA usage rate was 8%, and approximately 50% of SCA non-users were uninterested in trying SCAs. The most commonly used SCAs were NetDoktor and Ada. Surprisingly, SCAs were most frequently used in the age group of 51–55 years. No significant associations were found between SCA usage and satisfaction with the GP or the number of GP visits and SCA usage. Thematic analysis revealed skepticism regarding the results and recommendations of SCAs and discrepancies between users’ requirements and the features of apps.

Conclusion: SCAs are still widely unknown in the German population and have been sparsely used so far. Many participants were not interested in trying SCAs, and we found no positive or negative associations of SCAs and primary care.

Keywords

Symptom checker apps, primary care, self-triage, self-diagnosis, digital health, eHealth

Submission date: 15 March 2023; Acceptance date: 22 January 2024

[1]Institute of General Practice and Interprofessional Care, University Hospital Tübingen, Tübingen, Germany
[2]Institute of Software Engineering, University of Stuttgart, Stuttgart, Germany
[3]Institute of Applied Social Science, University of Applied Science Würzburg-Schweinfurt, Wurzburg, Germany
[4]Institute of Philosophy, University of Bremen, Bremen, Germany
[5]Institute of Occupational and Social Medicine and Health Services Research, University Hospital Tübingen, Tübingen, Germany
[6]Institute of Clinical Epidemiology and Applied Biometry, University Hospital Tübingen, Tübingen, Germany
[7]Faculty of Health Sciences, University of Potsdam, Potsdam, Germany
[8]Institute of Ethics and History of Medicine, University Hospital Tübingen, Tübingen, Germany

Corresponding author:
Anna-Jasmin Wetzel, Institute of General Practice and Interprofessional Care, University Hospital Tübingen, Osianderstr. 5, 72076 Tübingen, Germany.
Email: jasmin.wetzel@med.uni-tuebingen.de

Creative Commons CC BY: This article is distributed under the terms of the Creative Commons Attribution 4.0 License (https://creativecommons.org/licenses/by/4.0/) which permits any use, reproduction and distribution of the work without further permission provided the original work is attributed as specified on the SAGE and Open Access page (https://us.sagepub.com/en-us/nam/open-access-at-sage).

Introduction

Although the internet as a source of health information is ranked the lowest considering trust by patients compared to other sources such as the social environment or doctors, it is still frequently consulted.[1] Studies of health information seeking behaviour (HISB) indicate that 72% of Germans use the internet as a source of health information.[2] Further studies identified general practitioners (GPs) as the primary source of HISB in Germany, followed by medical specialists and pharmacists.[3]

HISB is defined by the type and quantity of health-related information individuals seek, the particular strategies employed to acquire information, and the sources they rely on.[4] This concept can be can be integrated into the broader framework of 'access to information', which is defined by the Universal Declaration of Human Rights as the entitlement 'to actively seek, receive and share information and ideas through any means, without regard to geographic boundaries'. This underscores the importance of recognising access to health information within this framework.[5]

Symptom checker apps (SCAs) are one example of a newly developed type of health app that can be used as a source in the context of HISB. SCAs are tools designed for medical laypeople and promise to offer self-triage as well as the classification of symptoms in a low-threshold setting. Users enter symptoms and receive an assessment of their conditions with plausible diagnosis suggestions as well as recommendations for further actions, such as seeing a doctor, going to an emergency room or staying at home and waiting.

As the relevance of these tools may increase due to an increasing scarcity of resources in the health care system such as lack of a sufficient number of GPs, adequate handling is important. In several countries, SCAs are also offered by federal organizations, such as NHS111[6] and the German National Association of Statutory Health Insurance.[7] Considering the reported lack of accuracy, it is important that SCA recommendations and the expected benefits resulting from their usage are based on solid evidence. SCAs are often claimed to be inaccurate and described as risk-averse, that is, they recommend seeking health care in non-urgent scenarios.[8–11] A recent study from 2022 based on the work of Semigran et al. (2015)[8] examined whether the triage performance of SCAs has improved in the last five years, and found that, on average, it has not.[12] In two SCA use cases, triage performance (advice on when emergency care is necessary and when no health care service is required) has decreased in the last 5 years.[12] The authors stated that triage capability still varies widely depending on the chosen app and the specific question—for example, the overall diagnostic accuracy differed between 25% and 80%.[12] A recent study found that SCAs recommended the use of primary care resources in 81% of the cases for which NICE (National Institute for Health and Care Excellence) clinical knowledge summaries indicate self-care.[9] Moreover, the safety of action recommendations decreases with urgency, which may result in an undersupply.[9] A recent study compared laypersons' triage accuracy with the accuracy of SCAs and found that the latter was not superior.[13] Despite the frequently insufficient accuracy[8–10,14] SCAs satisfy their users.[15,16] However, it remains unclear whether users are aware of their poor performance or if they continue to see the benefit in using SCAs despite the inaccuracy.

It is necessary to gain knowledge about SCA user groups, their recent extent of use and the consequences in the context of primary care. To date, the extent of SCA use and the reasons for their use are largely unknown.

To our knowledge, no previous studies have investigated the recent usage rate of SCAs or users' willingness or reluctance to use SCAs in Germany. Our study is part of the multi-disciplinary joint project CHECK.APP that investigates the ethical, legal and social implications of SCAs in Germany.[17] In our study, we conducted a survey to gain better knowledge about SCA use among the German population. Furthermore, we used user diaries and qualitative interviews to gain an in-depth understanding of user behaviour and potential implications for primary care. In this paper, we will present results from the survey and focus on the following research questions:

1. What is the degree of SCA utilization in Germany and which SCA software is the most used?
2. What role do SCAs play as part of the HISB?
3. Are there associations between SCA usage and the number of consultations of the GP and between usage and satisfaction with the GP present?
4. What are the reasons for the non-use of SCAs?

While questions 1–3 aim to provide a general overview about the usage of SCAs in Germany, the last research question aims to provide a more detailed understanding of the group of non-users, who have been hardly considered in empirical research.

Methods

Design

An explorative, cross-sectional survey was conducted using a mixed-methods design. Open-ended responses were assessed and processed using the qualitative data. The survey was available online and in a paper-and-pencil version.

Measurements

SCA usage was assessed in three categories: 'not an SCA user, and never heard of SCAs', 'not an SCA user, but

have already heard of SCAs' and 'SCA user'. We defined subjects as 'SCA users' if they stated that they had used SCAs at least once. SCA users were asked to report the software that they had already used in filter questions. SCA non-users were asked if they were interested in trying SCAs. Open-ended responses were assessed to explore intentional non-use.

Subjects were asked about the number of GP consultations per year on a five-level scale (less than once a year, once a year, every 3 months, once a month, and more than once a month). HISB was assessed with multiple-choice selection ('Assuming you have health complaints, where do you most commonly seek information?' Friends, books, doctors, Internet and SCAs). Satisfaction with GPs was assessed using an adapted version of the Questionnaire for Satisfaction in Ambulatory Care,[18] a validated four-item scale that assesses patient satisfaction with ambulatory care considering the concept of patient involvement. Answers on every item range between 0 and 3 and sum up to a score between 0 and 12. Permission to use the evaluated scales was requested individually via email and approved by the authors.

Data collection

Data were collected from November 2020 to June 2021. Different recruitment channels were used to reach a wide variety of participants and ensure that enough participants who used SCAs had a sufficient number of cases for statistical analysis. First, 50,000 German citizens were contacted via an e-mail to participate in the survey. The intended recipients were selected by an external partner (T + R Dialog Marketing (Berlin, Germany) and Acxiom (Neu-Isenburg, Germany)). Further participants were selected using e-mail lists at the University of Tübingen and the University Hospital of Tübingen, social media, and by cooperating with GPs.

Inclusion criteria for participation in the survey were the ability to give consent and German language skills of at least the B1 level of the Common European Framework of References for Languages.[19] The exclusion criterion was to have completed training as a medical doctor, as we assumed that the handling of SCAs by doctors differs from that of laypersons. In accordance with the ethical vote, all data were assessed anonymously, and participants provided written informed consent before starting the survey.

Quantitative analysis

Data processing and statistical analyses were conducted using R Version 4.1.1[20] and R Studio Version 1.4.[21]

Weighting. To maximise representativeness, data weighting based on summary statistics provided by the German Federal Statistical Office for the German population[22] was used. A raking procedure (*anesrake* package[23]) was applied to weigh the participants according to age, school education and gender. All analyses used these weights to reduce the bias associated with frequency deviations in the German population.

For descriptive data, frequencies (n) and percentages (%) were generated for categorical variables, and means and standard deviations (SD) were calculated. The median and interquartile range (IQR) were generated for continuous variables. To compare the three SCA (non-)user groups, the Kruskal–Wallis rank-sum test for complex survey samples was conducted for nominal- or ordinal-scaled variables.[24] To compare the HISB in different age groups with regard to SCA use, a weighted Pearson χ^2 test was performed. Due to the small number of missing data in the dataset and the assumptions of missing completely at random (MCAR), pairwise deletion was used in the case of missing single variables. Analyses that examined the associations between GPs and SCA use involved participants who stated that they had a GP.

Qualitative analysis

Data processing and qualitative analyses were conducted using MaxQDA Version 2020.[25] Open-ended responses regarding reasons for intentional non-use were analyzed using inductive thematic analysis following the principles outlined by Braun and Clark (2014).[26] The researchers had backgrounds in general practice (RK, MD, Dr, male), sociology (MK, Dr, male) and ethics (RM, Dr, female). Preliminary themes were discussed with members from other subprojects, and discrepancies that arose during theme synthesis were discussed with all authors and could be resolved by refining the themes. Data saturation was achieved in our study, as no further substantial aspects or new themes emerged during the data collection process. All authors approved the final themes.

Data integration

The quantitative and qualitative data were structured based on the thematic focus of the discussion section. This allowed merging[27] and higher-order integration of both quantitative and qualitative results were based on interpretations, which are described in the Discussion section.

Results

Respondents and sociodemographic characteristics

A total of 871 participants ($n = 116$, paper–pencil; $n = 755$ online) completed the survey. Of these, two had missing 5,6 values on SCA usage, and 19 were physicians and were therefore excluded; finally, 850 subjects were included in the analysis.

The sociodemographic characteristics of the participants are presented in Table 1. The three groups differed significantly in terms of age ($p = .01$) and educational level ($p = .03$). No group differences were observed with regard to gender, living area and chronic conditions, and being affiliated with a GP.

What is the degree of utilization of SCAs in Germany? Which SCA software is most used?

The SCA user group comprised 64 (7.5%) participants; the group of non-users who had already heard of SCAs comprised 213 (25.1%) participants, and non-users who had never heard of SCAs before were 573 (67.4%). Of the two non-user groups, 46% ($n = 362$) stated that they were interested in trying SCAs. Figure 1 shows the rate of usage stratified by interest in trying SCAs. Table 2 shows the most frequently used SCAs.

Which place do SCAs take as part of HISB?

The analysis of HISB stratified by age revealed significant group differences, with $\chi^2(44) = 138.06$, $p < .001$. Younger participants (18–30 years old) informed themselves more frequently via friends and less frequently via a doctor than those over 35 years of age. People over the age of 46 used books more frequently to inform themselves. SCAs were most frequently used by participants aged 51–55 years (Figure 2).

We regarded HISB as stratified for SCA user groups, and the proportions of health-seeking behaviour differed for the (non)-user groups of SCAs, with $\chi^2(8) = 355.12$, $p < .001$. One can observe that the proportion of participants who informed themselves via a doctor, friends and books was smaller in the SCA user group than in the other two groups. SCA users also consulted the Internet more frequently than both the non-user groups (Figure 3).

Are there associations between SCA usage and the number of consultations with the GP or the satisfaction with the GP?

Group differences in SCA usage and satisfaction with the GP were investigated and showed no significant differences, with $\chi^2(2) = 0.59$, $p = .74$. Moreover, no group differences were observed between SCA usage and the number of reported GP consultations and SCA usage, with $\chi^2(2) = 2.95$, $p = .23$ (Table 3).

Reasons for the intentional non-use of SCAs

Intentional non-users of SCA were asked, through an open-ended question, whether they were interested in trying SCAs, and if not, they were requested to give their reasons. We were able to generate several relevant and complex themes through a data analysis. The reflexive thematic analysis of the reasons for non-use yielded two overarching themes (Figure 4).

Trust in SCA vs. trust in health care providers. The first overarching theme comprised a comparison of trust in SCAs with trust in healthcare providers. In general, non-users expressed concerns about the trustworthiness of SCAs; in contrast, trust in the medical competence of established healthcare providers was higher. They indicated that healthcare providers are more capable of asking the 'right' questions about symptoms and consequently of making a correct diagnosis. They also preferred health care providers, who were perceived as an adequate initial point for health care and were able to offer physical examination as well as the opportunity to discuss findings. In addition, the trustworthiness of health care providers in ethical terms was also given more weight. Privacy and security concerns were expressed, which also influenced the perception of SCA's trustworthiness. From the perspective of non-users, SCAs are ethically less trustworthy than health care providers. Due to high trust in the established health sector, both in terms of medical and ethical competence, SCAs were assessed as irrelevant by non-users.

> *'I have more trust in the doctor than in an algorithm' (327)*

Barriers to SCA usage. The second overarching theme comprised the potential impact and causes of a discrepancy in the requirements of the user and the app. SCA apps were described as 'not *for everyone*'. Non-users saw themselves as unsuitable for SCA use due to their age, history, health conditions, lack of interest, lack of access to devices or lack of health literacy. Concerns were also expressed about the potential impact of SCA use, ranging from illness anxiety and fear of hypochondriasis to addiction.

> *I am a bit of a hypochondriac and I fear that if I used such devices I would think I was ill even more often than I do now. (240)*

On another theme, non-users indicated that they preferred other ways of seeking information on health topics:

> *I get enough information from friends and family about possible health problems. (178)*

Discussion

In our study, SCAs still seemed to play a minor role in HISB and primary care in Germany, with a usage rate of

Table 1. Sociodemographic characteristics of the three user groups.

Characteristics	Not An SCA User and Never Heard of SCAs, $N = 573$[a]	Not An SCA User, But Already Heard of SCAs, $N = 213$[a]	SCA User, $N = 64$[a]	P-value[b]
Age	52 (17)	48 (15)	42 (14)	.008
School education				.027
No graduation degree	10 (1.8%)	0 (0%)	0 (0%)	
Graduation after 9 years	198 (35%)	30 (14%)	10 (16%)	
Graduation after 10 years	173 (31%)	77 (36%)	28 (45%)	
A-levels	180 (32%)	106 (50%)	24 (39%)	
Gender				.2
Female	306 (54%)	93 (44%)	28 (45%)	
Male	258 (46%)	120 (56%)	35 (55%)	
Area				.4
Major city	47 (8.4%)	25 (12%)	4 (6.9%)	
Suburbs of a major city	31 (5.5%)	18 (8.3%)	5 (8.1%)	
Midsized city – town	327 (58%)	130 (61%)	31 (49%)	
Village	157 (28%)	40 (19%)	23 (36%)	
Chronic condition				.3
No	290 (52%)	102 (49%)	40 (65%)	
Yes	234 (42%)	99 (48%)	22 (35%)	
Prefer not to say	31 (5.6%)	7 (3.5%)	0 (0%)	
Personal GP	528 (94%)	193 (91%)	61 (94%)	.4

GP: general practitioner; SCA: symptom checker app.
[a]Mean (SD); *n* (%).
[b]Wilcoxon rank-sum test for complex survey samples; chi-squared test with Rao & Scott's second-order correction.

7.5%. The most commonly used SCA was the NetDoktor Symptom-Checker (45%), followed by Ada (30%). Surprisingly, the age group of 51–55 years reported the most frequent use of SCAs. As expected, SCA users informed themselves less frequently through doctors, books and friends.

Distribution and classification of SCAs in HISB

Our results considering the Internet research in the context of HISB were in line with other findings from Germany[1–3] and showed that the Internet research is the most used source in the context of HISB in addition to doctors' appointments. However, SCAs seem to play a minor role in HISB, as they had the lowest usage rate compared to the other assessed sources. The popularity and usage of SCAs in Germany are still subordinate to the use of health apps in general. In a German federal innovation analysis conducted in 2018, 54% of the participants stated that they were using health apps,[28] and the usage rate of these apps in general is consequently considerably higher than the low usage rate of SCAs.

We expected to observe a usage peak of SCA in the younger age group, as suggested by Meyer et al.;[29]

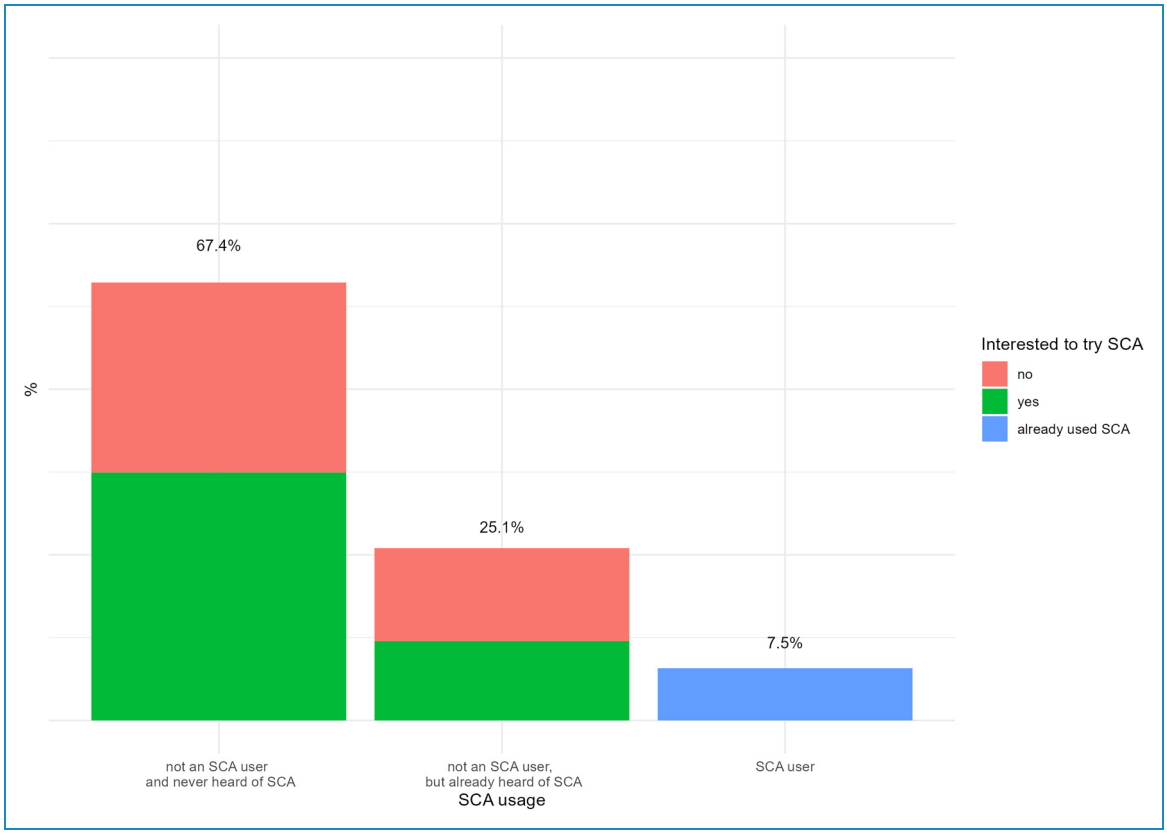


Figure 1. SCA usage and interest in trying SCAs. SCA: symptom checker app.

Table 2. Most used SCAs.

Most Used SCAs	$N = 64$[a]
NetDoktor SC	29 (45%)
Ada	20 (30%)
Other SCAs	19 (30%)
Diagnose Medizin app	3 (4.3%)
Symptomate	3 (4.3%)
WebMD	2 (2.9%)

SCA: symptom checker app.
[a]1n (%).

however, the age group that reported the highest SCA usage rate in our data was middle adulthood. One reason for the lower usage rate in younger age groups may be the preference for using general internet research as a competitor of SCAs. Further investigations are necessary to exclude the possibility of a sampling bias as a potential explanation for this finding.

Furthermore, we did not see older adults over 70 years of age using SCAs and only a minority of adults older than 60 years used SCAs. A study from 2018 reported that 16.5% of older adults in Germany used health apps and 37.5% used apps in general.[30] The lack of SCA utilization among certain demographics, such as individuals aged 55 and above, increases the likelihood of health inequality if the importance of SCAs continues to grow.

A further important aspect to consider in future research is whether different HISB sources exert an influence on one another, considering access to health information and/or access to health care in general. It might be particularly interesting to investigate SCA use in diverse health-care systems where access to primary care varies significantly.

Our study was carried out during peaks of the COVID-19 pandemic. Self-diagnosis is an important element in pandemic control, so it is likely that the general use of symptom checkers might increase during times of pandemic.[31] Various COVID-19 related symptom checkers had been developed to help citizens to self-diagnose a potential COVID-19 infection,[32] e.g., in Denmark, where a SCA was developed within days and was used 150,000 times in a few weeks.[33] In Germany, however, COVID-19 symptom checkers were not part of public pandemic management.

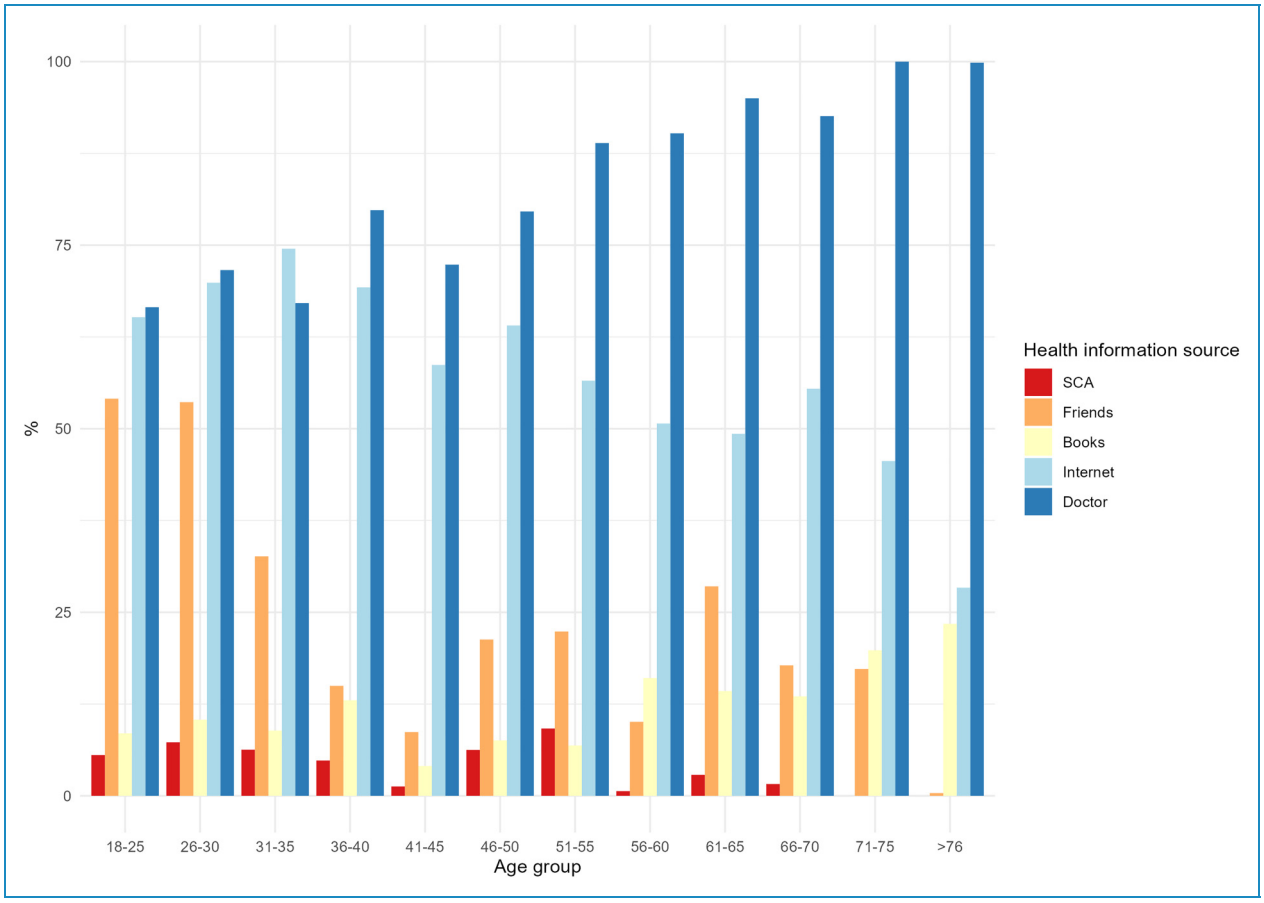


Figure 2. HISB stratified for the age group. HISB: health information seeking behaviour.

Associations between SCA use and GP care

The majority (75%) of SCA use comprised applications that aimed to process symptoms toward a probable diagnosis (such as NetDoktor Symptom-Checker and Ada), indicating a potential demand for SCAs as a tool in general practice.[11,34,35] A study from Switzerland in 2020 reported that out of 1000 participants, 546 had one or more health complaints within the previous two months – the majority of which were mainly dealt with by general practice. These results show the substantial strain that lies in the primary care health sector, which might be influenced by SCAs. However, the results of the present study indicate neither an association between SCA usage and GP consultations, nor an association between satisfaction with GPs and SCA use. The satisfaction rate with GPs in general was high in our sample, as it is usually for Germany[36] and Europe.[37] The association of satisfaction with GPs and SCA use may differ from that observed in our sample of other health systems with lower satisfaction rates.

In their study on the accuracy of SCAs, Ceney et al. (2021) concluded that SCAs may trigger additional resource utilization as it recommends, in over half of the cases, seeking health care, with primary care resources being suggested in approximately 81% of conditions where NICE clinical knowledge summaries indicate self-care.[9] A Danish study investigating the utilization of health data through new technologies discovered that, alongside data from wearables, the results obtained from SCAs are among the most mentioned for GPs during medical consultations for interpretation.[38] These findings may indicate an increasing relevance of interpreting symptom checker results in the context of primary care and are inconsistent with the promise of reduced health care utilization. Further studies have indicated risk-averse self-triage in the context of SCAs, in the sense that even rare diseases are not overlooked,[8,10] decreasing the accuracy of the action recommendations in non-urgent scenarios.[9] Schmieding et al. (2021) found that SCAs and laypersons performed the poorest in self-triage in non-urgency scenarios[13]; in contrast, another recent study reported a decrease in perceived urgency among users of the intended level of care after using a specific SCA.[39]

However, it remains unclear how laypersons handle especially non-urgent scenarios when using SCAs after receiving recommendations for action that suggest health care utilization. To prevent unnecessary utilization by

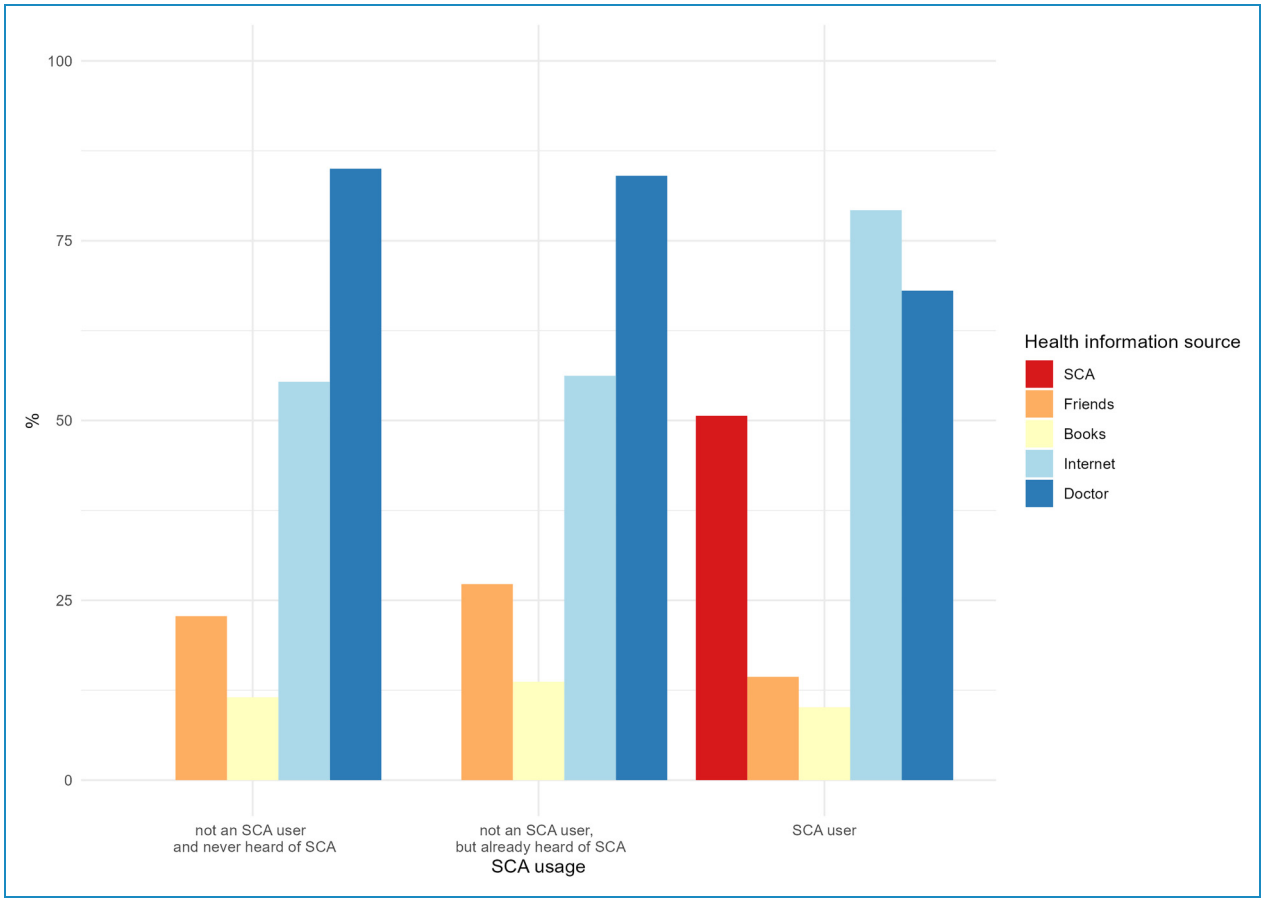


Figure 3. SCA use stratified for HISB. HISB: health information seeking behaviour; SCA: symptom checker app.

SCAs, it is necessary to understand the relationship between primary care utilization, handling of prior knowledge, action recommendations of SCAs and decision-making in laypersons. A study from 2022 that examined the impact of SCA explanations on laypersons' trust found that SCAs must tailor answers to each user's specific question and take the prior knowledge of the user into account.[40] When laypersons had prior knowledge about a disease, the answers given by SCA had little influence on their trust. However, these answers had a significant influence on trust when the need for information was higher.[41]

Reasons for intentional non-use of SCAs

One reason for the low usage rate was the lack of knowledge regarding the existence of SCAs (67.4%). In our sample, 25.1% stated that they knew about SCAs, but did not use it. Considering the non-user group that had already heard of SCAs, but had never used it, 47% expressed their disinterest in using SCAs. A qualitative analysis of the open-ended responses contemplating reasons for intentional non-use provided possible explanations. The general trustworthiness of SCAs was doubted, and it was perceived as an inadequate initial point for health care and unable to handle the complexity of diagnostic processes. Another reason was the perceived barriers to SCA use and limited benefits. Our results concerning privacy and security concerns align with other findings related to health apps in general.[42] Moreover, our findings regarding the usefulness of SCAs also corroborate the results from the literature, which reported a lack of perceived usefulness as hindering mobile app use in general.[43] The same study also reported a lack of app literacy as a further hindering factor, which coincides with our findings that suggest a mismatch of requirements between the user and SCAs.[43]

Strengths and limitations

The strengths of this study include the mixed-method and interdisciplinary team approaches; therefore, different perspectives and methods were integrated to make a broad spectrum of studies possible.

However, our results may not be applicable across countries, especially where healthcare systems, primary care provision, and cultural and social

Table 3. Variables considering GP care stratified for SCA use.

Characteristics	Not An SCA User and Never Heard of SCAs, $N = 573$[a]	Not An SCA User, But Already Heard of SCAs, $N = 213$[a]	SCA user, $N = 64$[a]	P-value[b]
ZAPA	10.00 (8.00, 12.00)	10.00 (8.00, 12.00)	10.00 (8.00, 11.00)	0.7
Number of GP visits				0.2
Less than once a year	45 (8.4%)	17 (8.9%)	1 (1.2%)	
Once a year	98 (18%)	61 (32%)	14 (23%)	
Once half a year	153 (29%)	45 (23%)	23 (37%)	
Every three months	178 (33%)	51 (26%)	22 (37%)	
Once a month	42 (7.8%)	11 (5.5%)	1 (1.4%)	
More than once a month	20 (3.7%)	8 (4.0%)	0 (0%)	

GP: general practitioner; SCA: symptom checker app.
[a]Median (IQR); *n* (%).
[b]Kruskal–Wallis rank-sum test for the complex survey samples.

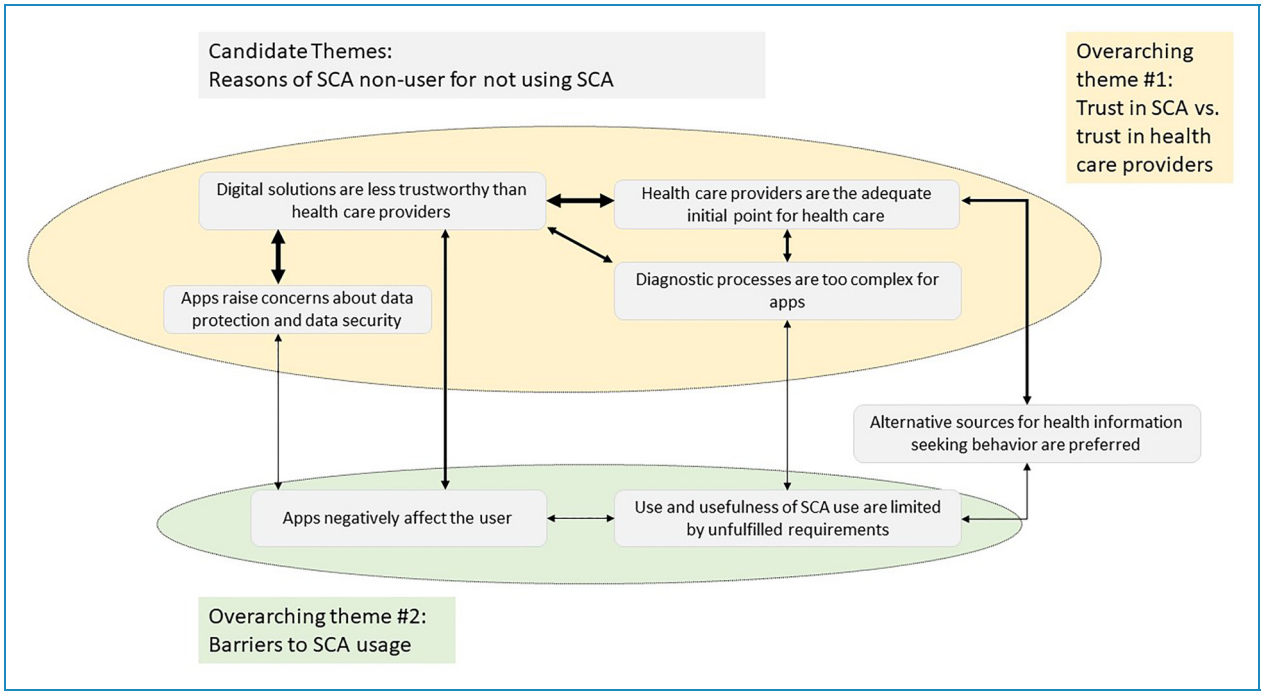


Figure 4. Themes of the thematical analysis considering the open-ended response to reasons for the intentional non-use of SCAs. SCA: symptom checker app.

conditions differ. The challenge of a heterogeneous wording about SCAs has previously been addressed and could not be solved completely in this study, too. As a consequence, we cannot exclude that some participants were unable to differentiate between a general Internet search and the use of SCAs, such as those accessed through a web browser. Furthermore, our data collection led to sampling bias toward younger and better-educated participants, we addressed this by weighting the data.

Conclusions

SCA use is still in its early stages. We observed a sample that was divided into skeptics and potential users, SCAs may have an important impact; however, our study does not indicate whether this impact is positive or negative. This is also a chance for regulations to prevent negative consequences and promote the benefits of these applications. Contrary to our expectations regarding the advertisement of SCAs, based on our study, SCAs are still widely unknown in the German population and sparsely used. Further research that examines primary care utilization in the context of SCA usage by taking into account a system perspective is necessary to identify the added value and possible unnecessary utilization or misallocation of resources through SCAs.

Acknowledgements: We would like to thank Michael Sand and Natalia Radionova for their assistance and guidance in this research. Furthermore, we would like to thank the participants of the cross-sectional survey. We acknowledge support from the Open Access Publication Fund of the University of Tübingen.

Declaration of conflicting interests: The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding: The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This work was funded by the German Federal Ministry of Education and Research (grant number 01GP1907A).

The work of the Institute of Occupational and Social Medicine and Health Services Research Tübingen was supported by an unrestricted grant from the Employerś Association of the Metal and Electric Industry Baden-Württemberg (Südwestmetall).

Ethical approval: Ethical approval for this study was obtained from the ethics committee of the University of Tübingen (ID: 464/2020BO).

Guarantor: SJ.

Contributorship: A-JW gained ethical approval for the study, researched the literature, conducted patient recruitment and processed as well as analysed the quantitative data. SJ, RK and A-JW conceived the study. SJ, RK, CP, MR, RR and H-JE were involved in the protocol development. NK supported patient recruitment and data processing. IR supported the weighting procedure. RM, MK and RK conducted the thematical analysis. A-JW wrote the first draft of the manuscript. All authors reviewed and edited the manuscript and approved the final version of the manuscript.

ORCID iDs: Anna-Jasmin Wetzel https://orcid.org/0000-0002-7127-5867
Monika Rieger https://orcid.org/0000-0002-7855-3663

References

1. Baumann E, Czerwinski F, Rosset M, et al. Wie informieren sich die menschen in Deutschland zum thema gesundheit? Erkenntnisse aus der ersten welle von HINTS Germany. *Bundesgesundheitsblatt–Gesundheitsforschung–Gesundheitsschutz* 2020; 63: 1151–1160.
2. Link E and Baumann E. Nutzung von gesundheitsinformationen im internet: personenbezogene und motivationale einflussfaktoren. *Bundesgesundheitsblatt - Gesundheitsforschung - Gesundheitsschutz* 2020; 63: 681–689.
3. Oedekoven M, Herrmann WJ, Ernsting C, et al. Patients' health literacy in relation to the preference for a general practitioner as the source of health information. *Bmc Fam Pract* 2019; 20: 1–10.
4. Lambert SD and Loiselle CG. Health information-seeking behavior. *Qual Health Res* 2007; 17: 1006–1019.
5. Assembly UG. Universal declaration of human rights. *UN Gen Assem* 1948; 302: 14–25.
6. NHS. 111 online - Get help for your symptoms https://111.nhs.uk/ (2023, accessed 14.03.2023).
7. Bundesvereinigung K. Patienten-Navi, https://www.116117.de/de/patienten-navi.php (2023).
8. Semigran HL, Linder JA, Gidengil C, et al. Evaluation of symptom checkers for self diagnosis and triage: audit study. *Br Med J* 2015; 351: h3480.
9. Ceney A, Tolond S, Glowinski A, et al. Accuracy of online symptom checkers and the potential impact on service utilisation. *Plos One* 2021; 16: e0254088. 20210715.
10. Hill MG, Sim M and Mills B. The quality of diagnosis and triage advice provided by free online symptom checkers and apps in Australia. *Med J Aust* 2020; 212: 514–519. 2020/05/12.
11. Radionova N, Og E, Wetzel AJ, et al. Impacts of symptom checkers for laypersons' self-diagnosis on physicians in primary care: scoping review. *J Med Internet Res* 2023; 25: e39219. 20230529.
12. Schmieding ML, Kopka M, Schmidt K, et al. Triage accuracy of symptom checker apps: 5-year follow-up evaluation. *J Med Internet Res* 2022; 24: e31810. 20220510.
13. Schmieding ML, Mörgeli R, Schmieding MAL, et al. Benchmarking triage capability of symptom checkers against that of medical laypersons: survey study. *J Med Internet Res* 2021; 23: e24475. 20210310.
14. Chambers D, Cantrell AJ, Johnson M, et al. Digital and online symptom checkers and health assessment/triage services for urgent health problems: systematic review. *BMJ Open* 2019; 9: e027743.
15. Knitza J, Mohn J, Bergmann C, et al. Accuracy, patient-perceived usability, and acceptance of two symptom checkers (Ada and rheport) in rheumatology: interim results from a randomized controlled crossover trial. *Arthritis Res Ther* 2021; 23: 112. 20210413.
16. Miller S, Gilbert S, Virani V, et al. Patients' utilization and perception of an artificial intelligence-based symptom assessment and advice technology in a British primary care waiting room: exploratory pilot study. *JMIR Hum Factors* 2020; 7: e19713. 20200710.
17. Wetzel AJ, Koch R, Preiser C, et al. Ethical, legal, and social implications of symptom checker apps in primary health care

(CHECK.APP): protocol for an interdisciplinary mixed methods study. *JMIR Res Protoc* 2022; 11: e34026. 20220516.
18. Scholl I, Hölzel L, Härter M, et al. Fragebogen zur zufriedenheit in der ambulanten versorgung – schwerpunkt patientenbeteiligung (ZAPA). *Klin Diagn und Eval* 2011; 4: 50–62.
19. Europe Co. *Common European Framework of Refrence for Languages: Learning, Teaching, Assessment*. Cambridge, UK: Cambridge University Press, 2001.
20. R Core Team. R: a language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing, 2016.
21. Team R. RStudio: Integrated Development for R. PBC, Boston: RStudio, 2021.
22. (Destatis) SB. Mirkozensus. 2019.
23. Pasek J and Pasek MJ. Package 'anesrake'. *The Comprehensive R Archive Network* 2018: 1–13.
24. Natarajan S, Lipsitz SR, Fitzmaurice GM, et al. An extension of the Wilcoxon rank sum test for complex sample survey data. *Journal of the Royal Statistical Society: Series C (Applied Statistics)* 2012; 61: 653–664.
25. VERBI Software. MAXQDA 2020. Berlin, Germany: VERBI Software, 2019.
26. Braun V and Clarke V. What can 'thematic analysis' offer health and wellbeing researchers? *Int J Qual Stud Health Well-Being* 2014; 9: 26152.
27. Fetters MD, Curry LA and Creswell JW. Achieving integration in mixed methods designs-principles and practices. *Health Serv Res* 2013; 48: 2134–2156.
28. Evers-Wölk M, Oertal B and Sonk M. Health apps. In: Bundestag Ootaat G (ed.) *TBA Fokus regarding report no 179*. Berlin: Büro für Technikfolgen-Abschätzung beim Deutschen Bundestag, 2018, pp.1–4.
29. Meyer AND, Giardina TD, Spitzmueller C, et al. Patient perspectives on the usefulness of an artificial intelligence-assisted symptom checker: cross-sectional survey study. *J Med Internet Res* 2020; 22: e14679. 2020/02/06.
30. Rasche P, Wille M, Bröhl C, et al. Prevalence of health app use among older adults in Germany: national survey. *JMIR Mhealth Uhealth* 2018; 6: e26.
31. Judson TJ, Odisho AY, Neinstein AB, et al. Rapid design and implementation of an integrated patient self-triage and self-scheduling tool for COVID-19. *J Am Med Inform Assoc* 2020; 27: 860–866.
32. Lai L, Wittbold KA, Dadabhoy FZ, et al. Digital triage: novel strategies for population health management in response to the COVID-19 pandemic. In: *Healthcare*. Amsterdam, the Netherlands: Elsevier, 2020, 100493.
33. Haase CB, Bearman M, Brodersen J, et al. 'You should see a doctor', said the robot: reflections on a digital diagnostic device in a pandemic age. *Scand J Public Health* 2021; 49: 33–36.
34. Hajek A, Bock J-O and König H-H. Which factors affect health care use among older Germans? Results of the German ageing survey. *BMC Health Serv Res* 2017; 17: 1–8.
35. Busse R and Stahl J. Integrated care experiences and outcomes in Germany, The Netherlands, and England. *Health Aff* 2014; 33: 1549–1558.
36. Gottfried M. Forsa-Umfrage - GKV-Versicherte sind mit der medizinischen Versorgung zufrieden, (2018, accessed 07.12.2022).
37. Detollenaere J, Hanssens L, Schäfer W, et al. Can you recommend me a good GP? Describing social differences in patient satisfaction within 31 countries. *Int J Qual Health Care* 2018; 30: 9–15.
38. Haase CB, Ajjawi R, Bearman M, et al. Data as symptom: Doctors' responses to patient-provided data in general practice. *Soc Stud Sci* 2023; 53: 522–544.
39. Winn AN, Somai M, Fergestrom N, et al. Association of use of online symptom checkers with patients' plans for seeking care. *JAMA Netw Open* 2019; 2: e1918561.
40. Kopka M, Schmieding ML, Rieger T, et al. Determinants of laypersons' trust in medical decision aids: randomized controlled trial. *JMIR Human Factors* 2022; 9: e35219.
41. Woodcock C, Mittelstadt B, Busbridge D, et al. The impact of explanations on layperson trust in artificial intelligence-driven symptom checker apps: experimental study. *J Med Internet Res* 2021; 23: e29386. 20211103.
42. Zhou L, Bao J, Watzlaf V, et al. Barriers to and facilitators of the use of mobile health apps from a security perspective: mixed-methods study. *JMIR Mhealth Uhealth* 2019; 7: e11223. Original Paper 16.04.2019.
43. Peng W, Kanthawala S, Yuan S, et al. A qualitative study of user perceptions of mobile health apps. *BMC Public Health* 2016; 16: 1–11.

2.3. Publication 3

Publication 3, titled "Only the Anxious Ones? Identifying Characteristics of Symptom Checker App Users: A Cross-sectional Survey," identifies key predictors of SCA usage based on user characteristics. This study was published in BMC Medical Informatics and Decision Making (Impact Factor 2022: 3.5).

BMC Medical Informatics and Decision Making

RESEARCH

Open Access

Only the anxious ones? Identifying characteristics of symptom checker app users: a cross-sectional survey

Anna-Jasmin Wetzel[1*], Malte Klemmt[2], Regina Müller[3], Monika A. Rieger[4], Stefanie Joos[1] and Roland Koch[1]

Abstract

Background Symptom checker applications (SCAs) may help laypeople classify their symptoms and receive recommendations on medically appropriate actions. Further research is necessary to estimate the influence of user characteristics, attitudes and (e)health-related competencies.

Objective The objective of this study is to identify meaningful predictors for SCA use considering user characteristics.

Methods An explorative cross-sectional survey was conducted to investigate German citizens' demographics, eHealth literacy, hypochondria, self-efficacy, and affinity for technology using German language–validated questionnaires. A total of 869 participants were eligible for inclusion in the study. As $n = 67$ SCA users were assessed and matched 1:1 with non-users, a sample of $n = 134$ participants were assessed in the main analysis. A four-step analysis was conducted involving explorative predictor selection, model comparisons, and parameter estimates for selected predictors, including sensitivity and post hoc analyses.

Results Hypochondria and self-efficacy were identified as meaningful predictors of SCA use. Hypochondria showed a consistent and significant effect across all analyses OR: 1.24–1.26 (95% CI: 1.1–1.4). Self-efficacy OR: 0.64–0.93 (95% CI: 0.3–1.4) showed inconsistent and nonsignificant results, leaving its role in SCA use unclear. Over half of the SCA users in our sample met the classification for hypochondria (cut-off on the WI of 5).

Conclusions Hypochondria has emerged as a significant predictor of SCA use with a consistently stable effect, yet according to the literature, individuals with this trait may be less likely to benefit from SCA despite their greater likelihood of using it. These users could be further unsettled by risk-averse triage and unlikely but serious diagnosis suggestions.

Trial Registration The study was registered in the German Clinical Trials Register (DRKS) DRKS00022465, DERR1-https://doi.org/10.2196/34026.

Keywords Digital health, Mental health, eHealth, mHealth, Cyberchondria, Self-diagnosis, Symptom Checker, Patient safety

*Correspondence:
Anna-Jasmin Wetzel
jasmin.wetzel@med.uni-tuebingen.de

Full list of author information is available at the end of the article

© The Author(s) 2024. **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/. The Creative Commons Public Domain Dedication waiver (http://creativecommons.org/publicdomain/zero/1.0/) applies to the data made available in this article, unless otherwise stated in a credit line to the data.

Introduction

Symptom checker Apps (SCAs) are eHealth applications designed to support laypeople in assessing their symptoms and receiving recommendations on medically appropriate actions related to their health [1]. Users can input their health-related information into SCAs through a chatbot or search strings, and SCAs retrieve and categorize the input. Some SCAs are advertised as AI-based, and most generate healthcare-related information and recommendations for actions based on user input [2].

Although SCAs are already in use, their impact on healthcare systems remains poorly understood. Recent scoping reviews described ambiguous effects of SCAs [1, 3], indicating that they could both reduce or induce oversupply. The effectiveness of SCAs in delivering adequate and precise information and recommendations must be considered. Additionally, the possible impact of SCAs on healthcare systems depends on several factors, including the characteristics of SCAs and how SCAs are used. Finally, the impact of SCAs on users' health-related behavior, such as seeking healthcare, must also be considered.

Recent studies have shown that the diagnostic accuracy and triage capabilities of SCAs are highly variable. A recent study reported a triage accuracy for primary conditions varying between 48.8% and 90.1% [1]. Additionally, a significant disparity in diagnostic accuracy between SCAs and emergency physicians has been reported. While SCAs correctly identified the primary diagnosis in only 30% of cases, emergency physicians achieved a much higher accuracy rate, successfully diagnosing 81% of cases [4]. In addition, another study found that medical laypeople still outperformed SCAs [5]. Consequently, SCAs currently struggle to reliably assist patients in navigating healthcare and addressing adequate medical recommendations.

Understanding the impact of SCAs on the healthcare system requires consideration of user demographics, such as (e)health literacy and attitudes toward technology [3, 6, 7]. Research indicates that SCA users are often female, well-educated, Caucasian, with health insurance and a regular healthcare provider [8, 9]. Recent studies showed that health literacy levels in Germany have declined over the years with the reported uncertainty being mainly related to online resources [10]. However, some users found SCAs useful for self-diagnosis and reported positive health effects [11], while others had problems giving and interpreting concrete information on symptom time patterns or severity [12]. Such difficulties may initiate unnecessary healthcare-seeking behavior, although the evidence remains inconclusive [13]. Additionally, increased eHealth literacy may lead to greater subjective trust in SCAs and the ability to critically evaluate their recommendations, but not necessarily to a change in actual trust-based behavior [14]. Lastly, user attitudes toward technology play a significant role, with "tech seekers" being more likely to use SCAs in the future compared to "tech rejectors" and "unsure acceptors" [15]. Concurrently with internet research, the usage of SCAs may also magnify preexisting user characteristics associated with unwarranted healthcare-seeking tendencies rather than operating independently [16]. As an example, SCA may worsen hypochondria, similarly to how internet research is already known to do among vulnerable patient groups [16].

There is a research gap concerning the influence of concepts such as hypochondria, self-efficacy, technology affinity, and health literacy on the use of SCAs. Therefore, the aim of this explorative study was to identify meaningful predictors for SCA use considering user characteristics.

Methods

An explorative cross-sectional survey was conducted. The survey was available online or as a paper and pencil version. The STROBE (Strengthening the Reporting of Observational Studies in Epidemiology) checklist [17] was applied.

Measurements

Due to the limited literature on SCAs, pilot interviews with SCA users and SCA experts were conducted to ensure a meaningful concept selection for the survey content. In addition, to identify potential characteristics of the user group, we drew on literature related to the use of health applications.

Thus, the following concepts were selected: eHealth literacy [18], hypochondria [19], self-efficacy [20], and affinity for technology [21]. Table 1 presents a comprehensive overview, detailing the reliability, validity, scale, and scoring of the evaluated scales (General Life Satisfaction Short Scale [22], German Version of the eHealth Literacy Scale [23], Whiteley Index [24, 25], General self-efficacy short scale [26], Ultra-Short Scale for Assessing Affinity for Technology Interaction [27]) used in this study.

Furthermore, the presence of chronic diseases, private screen time (as a potential indicator of smartphone use) were assessed. Sociodemographic variables such as age, gender, and school education were also assessed in this study.

Recruitment

The survey was conducted from November 2020 to June 2021. The sample comprised different recruiting strands to reach a wide variety of participants and ensure a sufficient number of SCA users for the statistical analysis. In the first strand, $n=50.000$ German citizens were

Table 1 Overview, detailing the reliability, validity, scale, and scoring of the evaluated scales used in this study

Concept	Title of the Instument (Abbreviation)	Reliability	Validity	Number of Items	Scale	Score (Range)
Life satisfaction	General Life Satisfaction Short Scale (23) (L1)	retest = 0.82 (23)	evidence for construct validity (23)	1	10 point Likert	Sum Score (0–10)
eHealth literacy	German Version of the eHealth Literacy Scale (24) (G-eHeals)	Cronbach's α = 0.83-0.88 (24)	construct validity was supported by correlation patterns with different scales and constructs (24)	8	5 point Likert	Sum Score (0–40)
Hypochondria	Whiteley Index (25) (WI)	Cronbach's α = 0.76–0.8 (26)	high associations with health anxiety and illness behavior subscales (26)	14	binary („yes" / „no")	Sum Score (0–14)
Self-efficacy	General self-efficacy short scale (27) (ASKU-S)	McDonalds ω = 0.81 bis 0.87 (27)	Indications of content, factorial, convergent, discriminant and predictive validity (27)	3	5 point Likert	Mean Score (1–5)
Technic affinity	Ultra-Short Scale for Assessing Affinity for Technology Interaction (28) (ATI-S)	McDonalds ω = >0.8 (28)	sufficiently valid assessment (28)	4	6 point Likert	Mean Score (1–6)

contacted via mail to participate in the survey. The intended recipients were representatively selected by an external partner (T + R Dialog Marketing (Berlin, Germany) and Acxiom (Neu-Isenburg, Germany)). Further participants were recruited by mailing lists of the University of Tübingen and the University Hospital of Tübingen, social media and by cooperating GP practices. The second strand aimed to reach SCA users only; therefore, participants were only included if they had SCA experience. Targeted advertisements via social media, the social channels of the University Hospital of Tübingen, the homepage of a German newspaper and the social channels of federal health insurance were conducted to recruit further SCA users.

Data exclusion

We assumed a missing completely at random mechanism (only single values were unplausible or missing, omitted by chance). Participants with missing data on the primary outcome were excluded ($n = 2$). Furthermore, physicians ($n = 19$) were excluded due to the assumption that their medical knowledge would have a significant influence on SCA usage.

Statistical analysis

The primary outcome variable was whether participants had already used SCAs. Statistical analyses were conducted in different steps. The first step comprised variable selection using a least absolute shrinkage and selection operator (LASSO [28]) regularized logistic regression analysis considering nine predictors (as listed in Table 2). The second-step model comparison involved an intercept-only model, a full model and a model with the selected predictor using conventional logistic regression. In the third step, we utilized the identified predictors to derive parameter estimates and p-values. This process led to the determination of the main analysis parameters. A post-hoc analysis was conducted in the fourth step.

Propensity score matching

Users and non-users were matched with propensity score matching [29, 30] on an initial set of potential confounders [31]. Confounder covariates included school education and age, as we assumed that we reached a younger and better-educated user population due to our targeted recruiting strategy via social media and university mailing lists. A nearest neighbor matching algorithm [29] was applied. Missing data on the predictors were imputed using a random forest approach [32] that enables the imputation of missing information in mixed-data (categorical and continuous). Out-of-bag errors were considered [32].

Predictor selection using LASSO regularized logistic regression

The participants were divided into training (70%) and test (30%) data sets. The training data set was used to fit a model on the given data, and the test data set was used to evaluate the model [33]. A 0.632 bootstrap estimator [34] was applied as the resampling method for lambda selection. The sensitivity, specificity, Positive Predictive Value (PPV), Negative Predictive Value (NPV) and accuracy rate were calculated by fitting to the test data set. An overview of the predictors included can be found in Table 2.

Table 2 Overview of the potential predictors stratified for SCA use and univariate analysis

Predictors	Non-user, $N=67$[1]	User, $N=67$[1]	*p*-value[2]
Gender			0.15
Female	49 (73%)	41 (61%)	
Male	18 (27%)	26 (39%)	
Chronic Disease	22 (33%)	31 (46%)	0.12
Subj. rated health	1.9 (0.8)	2.2 (0.7)	0.018
Life satisfaction (L1)	7.4 (1.6)	6.9 (1.5)	0.074
eHealth literacy (G-eHeals)	30.0 (5.6)	30.7 (5.4)	0.5
Hypochondria (WI)	2.6 (2.8)	4.8 (3.2)	< 0.001
Self efficacy (ASKU-S)	4.1 (0.6)	3.8 (0.7)	0.040
Technic affinity (ATI-S)	3.5 (1.2)	3.7 (1.1)	0.5
Private screen time			0.2
0–30 min	3 (4.5%)	2 (3.0%)	
30–60 min	14 (21%)	10 (15%)	
1–2 h	19 (28%)	19 (28%)	
2–3 h	15 (22%)	9 (13%)	
3–4 h	11 (16%)	13 (19%)	
> 4 h	5 (7.5%)	14 (21%)	

[1]n (%); Mean (Standard Deviation)

[2]chi-squared test with Rao & Scott's second-order correction; Wilcoxon rank-sum test for complex survey samples

Model comparison of an intercept only, a full model and a model with the identified predictor

A conventional logistic regression was fit on the complete matched data set to derive odds ratios (ORs) and confidence intervals (CIs). To identify potential multicollinearity, we employed the Variance Inflation Factor (VIF), which assesses the variance of a coefficient within the full model in comparison to its variance when modeled independently [33]. A VIF value exceeding 5 were considered as indicative of significant collinearity [33]. The Akaike information criterion (AIC) of a full model, an intercept-only model and the model with the LASSO selected predictors was compared to assess model performance. The smaller the AIC, the better the performance of the model [33].

Parameter estimators, CI and *p*-values of the models, including the selected predictors

Parameter estimators were derived from conventional logistic regression. Two sensitivity analyses were conducted to ensure the robustness of the ORs, CIs and *p* values considering different sample compositions. We applied a different matching algorithm (full optimal matching [29]) for the sensitivity analysis. Additionally, we used the whole sample without matching for sensitivity in the second analysis.

Post hoc Analysis: categorization of the WI

Finally, a post hoc analysis was conducted considering a predictor identified in step 3. The variable was dichotomized to identify clinically relevant persons, and a Pearson's χ^2 test was conducted.

Data Processing

Data processing and statistical analyses were conducted with R Version 4.1.1 [35, 36] and R Studio Version 1.4 [37].

Results

A total of 869 participants ($n=116$ paper-pencil, $n=753$ online) completed the survey. As participants were matched 1:1 and 67 users finished the survey, the final analysis included $n=134$ participants. The median age of the population was 31 (IQR 24–49), and 67% were female. The matched variables (age and school education) were well balanced between the user and non-user groups (Love Plot Supplemental Fig. 1).

Table 1 describes the matched sample stratified for SCA use, including all predictors used in the LASSO regression. Univariate analyses were conducted for all predictors. In addition to subjective rated health, hypochondria and self-efficacy showed a significant association with SCA use.

Identification of meaningful predictors

The training data set comprised 93 participants. The test data set comprised 41 participants. Nine variables were initially considered for predictor selection, as detailed in Table 2. The selection process, which involved a LASSO regularized logistic regression, identified two variables with nonzero coefficients: hypochondria (WI) and self-efficacy (ASKU). Consequently, these two variables were chosen as predictors in the conventional logistic regression model. The LASSO coefficient profiles against log (λ) are shown in Fig. 1, as is the bootstrapped ROC curve for the regularization parameter λ. Figure 2 shows the

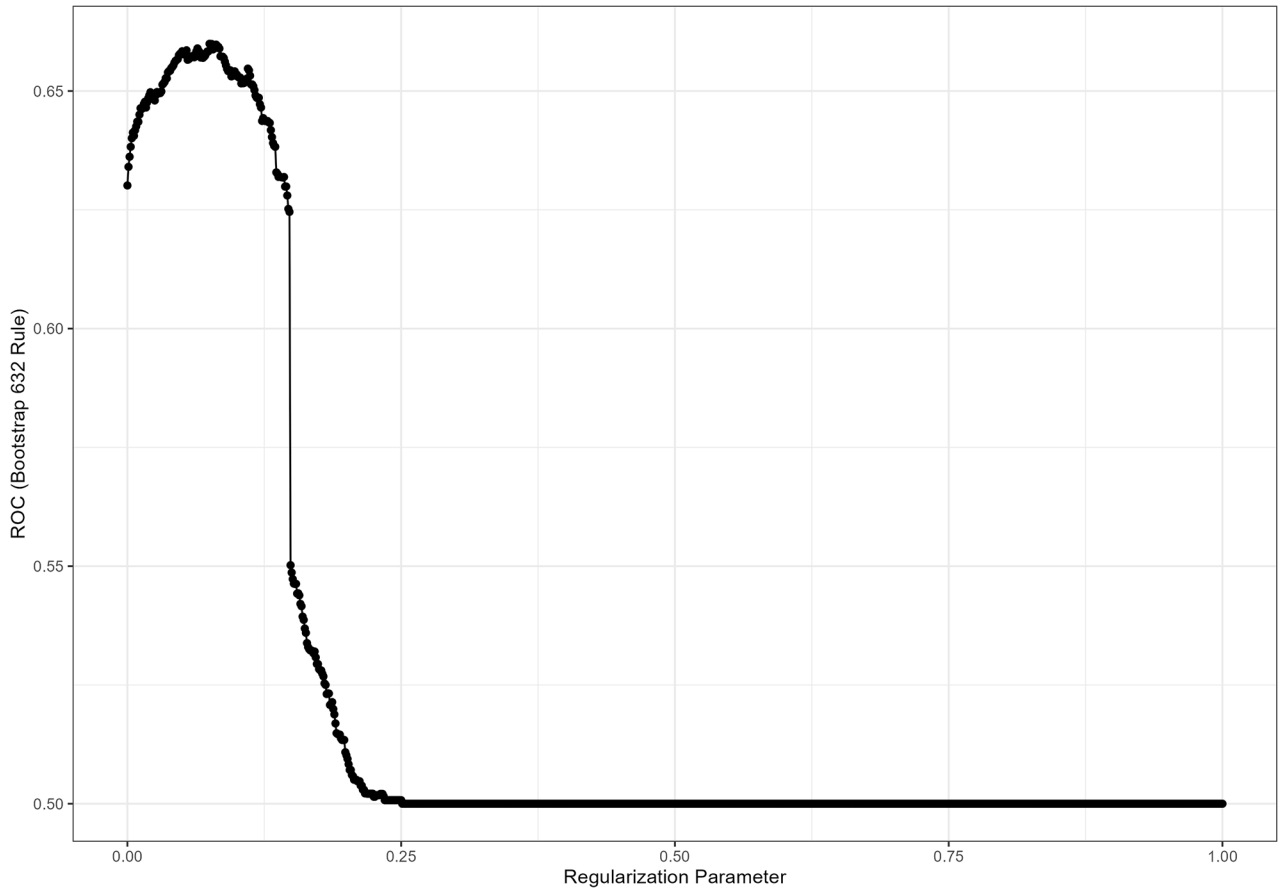


Fig. 1 Bootstrapped ROC curve for λ

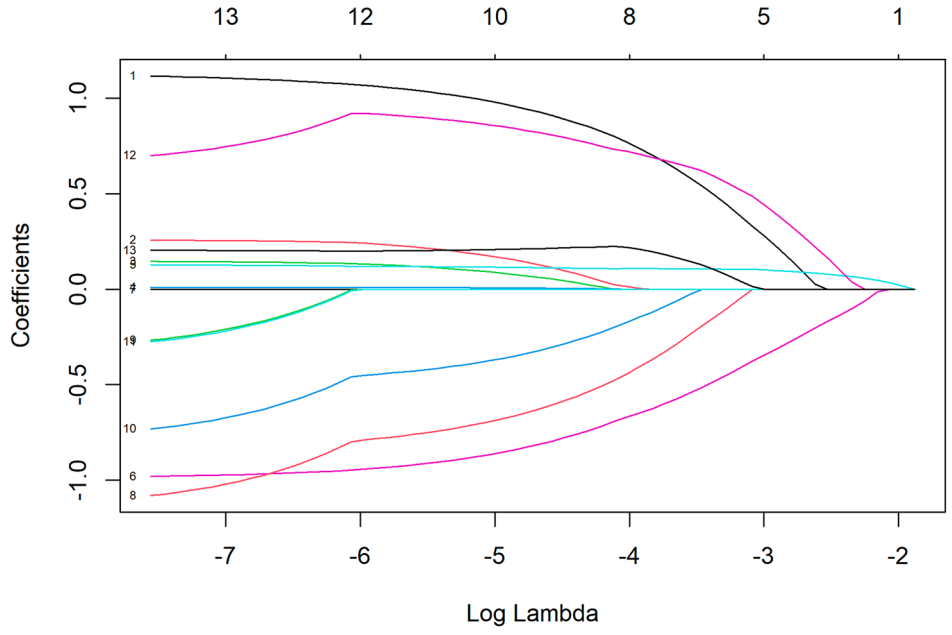


Fig. 2 LASSO coefficient profiles against log (λ), Lambda = 0.112 when the error of the model is minimized, and 2 variables were selected

Table 3 Model evaluation of the LASSO regression of the test data set

Sensitivity (%)	Specificity (%)	NPV (%)	PPV (%)	Balanced Accuracy
75.0	81.0	77.3	79.0	78.0

Table 4 Results of the conducted logistic regression and the sensitivity analysis

	Resulting logistic regression			Sensitivity analysis 1			Sensitivity analysis 2		
Characteristic	**OR[1]**	**95% CI[1]**	***p*-value**	**OR[1]**	**95% CI[1]**	***p*-value**	**OR[1]**	**95% CI[1]**	***p*-value**
Hypochondria (WI)	1.26	1.11, 1.46	< 0.001	1.24	1.14, 1.35	< 0.001	1.24	1.14, 1.35	< 0.001
Self efficacy (ASKU)	0.64	0.34, 1.15	0.14	0.93	0.63, 1.39	0.7	0.93	0.63, 1.39	0.7

[1]OR = Odds Ratio, CI = Confidence Interval

Table 5 Post hoc analysis with categorized WI stratified for the user group

Hypochondria Scale (WI)	Non-user, *N* = 67[1]	User, *N* = 67[1]	*p*-value[2]
Continuous	2.0 (1.0, 3.5)	5.0 (2.0, 7.0)	< 0.001
Categorized with a cut-off > 5			< 0.001
Hypochondria	8 (12%)	35 (52%)	
No hypochondria	59 (88%)	32 (48%)	

[1]Median (IQR); n (%)

[2]Wilcoxon rank sum test; Pearson's Chi-squared test

LASSO coefficient profiles against log (λ), Lambda = 0.112 when the error of the model is minimized, and 2 variables were selected. Sensitivity, specificity, Negative / Positive Predictive Values (NPV / PPV) and balanced accuracy can be found in Table 3.

Model comparisons

The AIC of the full model was 184.43, and the intercept-only model derived an AIC of 187.76. The logistic regression model based on the results of the LASSO variable selection (WI and ASKU-S) had the lowest AIC of 172.21 and therefore showed an improved performance compared to the full and intercept-only model. The VIF = 1.035 showed no considerable multicollinearity.

Parameter estimators and predictor robustness

Table 4 shows the odds ratios, confidence intervals of the odds ratios, and *p* values of the logistic regression and its sensitivity analyses comprising the previously identified predictors, hypochondria and self-efficacy. The OR of the predictor hypochondria (WI) showed a similar value (1.24–1.26) for all three models and a significant *p* value ($P < .001$). The OR size corresponds to a small effect [38]. The ORs for self-efficacy, measured across the three models, were not statistically significant ($P > .05$). Additionally, the variation inflation factor for all models was low (VIF < 1.06), indicating no considerable multicollinearity.

Post hoc analysis

Pearson's χ^2 test revealed a significant difference between non-users and users among participants with clinically relevant levels of hypochondria on the WI.

Over half of the SCA users had a WI sum score higher than the cut-off of five, indicating clinically relevant hypochondria (Table 5).

Discussion

In this exploratory study, we identified WI-assessed hypochondria as a reliable predictor for SCA use. This predictor consistently affected all analyses, including the two sensitivity analyses. Furthermore, lower values of self-efficacy assessed with the ASKU-S were identified as a positive predictor for SCA use in the main analysis. The sensitivity analyses did not replicate the effect of this variable; thus, its role remains unclear due to the rather moderate sample size.

Comparison with prior work

Hypochondria was identified as a predictor for SCA use and revealed a stable effect throughout our analyses. Over half of the SCA users had a WI sum score higher than the cut-off of five, indicating clinically relevant hypochondria (Table 4). This level of anxiety may affect a patient's ability to adequately handle action recommendations and symptom classifications. Thus, these SCA users might be susceptible to the negative effects of SCA use. Hypochondria in the context of SCAs can be classified as cyberchondria, considering the working definition of Vismara [39]. A 2020 study discouraged self-diagnosis using SCAs among cyberchondriac patients and emphasized adjusting expectations accordingly when accessing health information online [40]. Another recent study revealed that some people with high WI (hypochondria) scores felt worse after online symptom checking, while others with low scores felt better [41]. Given this literature and our findings, it appears that patients with health

anxiety are less likely to benefit from SCAs, despite being more inclined to use them. The transferability of results from online health-related use to SCAs is important to consider, as they suggest that prolonged use is associated with increased functional impairment and anxiety both before and after checking [41]. The impact of using SCAs on health-anxious patients remains unclear and warrants further investigation.

Furthermore, we examined self-efficacy as a meaningful predictor since a recent study indicated an association between self-efficacy and the adoption of SCA use [20]. The results of the predictor self-efficacy were ambiguous with differing effect sizes in our analyses. It is still uncertain how much self-efficacy contributes to determining the usage of SCAs.

Affinity for technology was another variable we considered since the literature indicated a potential association [15]. A study that examined SCA user profiles with a latent class analysis revealed that the latent class of "tech seekers" showed the highest odds of using SCA [15]. However, the results in our rather moderate sample do not suggest an association between affinity for technology and SCA use. Reasons for the discrepancy might be the different operationalization of the concepts (e.g., a scale rather than profiles) or the different study populations.

The broad use of SCAs can lead to individual- and systemic-level effects. SCAs could lead to a misuse of health care resources [4], such as users visiting emergency departments too early or too often. As a result, these users in nonurgent conditions put further strain on the health system by possibly increasing costs and taking resources from patients who need emergency care [3, 42]. To mitigate these risks, software developers should provide transparent information about the potential dangers of using SCAs. This information could be presented in the form of an instruction leaflet, available after downloading or using an SCA in a browser. The instruction should clearly state that using SCAs may increase health anxiety. The language used in the instruction should be concise and easy to understand so that users can absorb the information and take appropriate action.

The existing knowledge about SCAs should be used to improve SCA design in the best possible way and implement improvements to minimize the negative effects and strengthen the potential positive effects. Physicians should be trained to consider pre-informed patients and promote dialog. It is necessary to better understand the relationships between cyberchondria, hypochondria, and e-health literacy in the context of SCA use to derive recommendations for systemic interventions and plan targeted and helpful interventions.

Strengths and limitations

In this study, we conducted an industry-independent investigation of SCA users. Furthermore, our research did not limit itself to a single SCA application; instead, we examined the usage patterns across various types of SCA applications, enhancing the generalizability of our findings. Additionally, by matching users and non-users based on age and education, we controlled for these variables, thereby strengthening the reliability of our analysis.

A limitation of this study is the cross-sectional design we employed. This approach restricts our ability to infer causal relationships between variables, as it only provides a snapshot in time, thereby limiting our understanding of the dynamics and directionality of the relationships observed. Additionally, the recruitment of this study lead to younger and better-educated individuals introduces a potential selection bias. Our study's moderate sample size, while adequate for exploratory purposes, may not capture the full spectrum of SCA usage characteristics. Conducting this research on a larger scale would be beneficial to validate our findings and identify more nuanced predictors of SCA use. Moreover, our approach of double targeting SCA users to ensure a higher response rate might lead to response bias, potentially resulting in an overrepresentation of the views and behaviors of more engaged or interested users. In light of these limitations, future research should consider longitudinal studies involving more diverse and larger samples.

Conclusions

Hypochondria emerged as a significant predictor of SCA use in our sample, with a consistently stable effect. Over half of the SCA users had clinically relevant hypochondria considering their values on the WI, which may impact their ability to handle SCAs effectively. According to the literature, persons with hypochondria are less likely to benefit from SCA. These users could be further unsettled by risk-averse triage and unlikely but serious diagnosis suggestions. Software developers should provide transparent information about the potential dangers of using SCAs, including that SCA use may increase health anxiety. Individuals with higher levels of health anxiety (hypochondria) might experience increased anxiety or functional impairment due to SCA use. Users should be cautious of over-relying on SCAs for health information and diagnosis. For healthcare professionals, training in addressing patient concerns arising from SCA use may be beneficial, particularly for managing individuals with high health anxiety. Further, the widespread use of SCAs may potentially lead to the misuse of healthcare resources, with nonurgent cases increasing the burden on emergency services.

Abbreviations

SCAs	Symptom Checker Apps
STROBE	Strengthening the Reporting of Observational Studies in Epidemiology
DRKS	German Clinical Trials Register
L1	General Life Satisfaction Short Scale
G-eHeals	German Version of the eHealth Literacy Scale
WI	Whiteley Index
ASKU-S	General self-efficacy short scale
ATI-S	Ultra-Short Scale for Assessing Affinity for Technology Interaction
LASSO	Least Absolute Shrinkage and Selection Operator
VIF	Variance Inflation Factor
PPV	Positive Predictive Value
NPV	Negative Predictive Value
OR	Odds Ratio
CI	Confidence Interval

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s12911-024-02430-5.

Supplementary Material 1

Supplementary Material 2

Supplementary Material 3

Acknowledgements

We thank Maximilian Pilz and Fabian H. Klopfer for their advice. Further we would like to thank the participants of the cross-sectional survey. We acknowledge the assistance of OpenAI's ChatGPT language model for providing support with the revision of certain sections of this manuscript.

Author contributions

JW, RK and SJ designed the study and the study materials. JW collected the data, and performed the analyses. RK verified the data and contributed to the interpretation of the results. JW also wrote the first draft of the manuscript, with critical input from RK, SJ, MR, MK, RM. All authors reviewed and approved the final manuscript.

Funding

This study was funded by Federal Ministry of Education and Research, abbreviated BMBF, is a cabinet-level ministry of Germany [grant number: 01GP1907A]. The funder played no role in the study design, data collection, data analysis and interpretation, or manuscript writing.
The work of the Institute of Occupational and Social Medicine and Health Services Research Tübingen is supported by an unrestricted grant from the employers´ association of the metal and electric industry Baden-Württemberg (Südwestmetall).
Open Access funding enabled and organized by Projekt DEAL.

Data availability

The datasets used and/or analysed during the current study are available from the corresponding author on reasonable request.

Declarations

Ethics approval and consent to participate

Ethical approval for this study was obtained from the ethics committee of the University of Tübingen (ID: 464/2020BO). All data was assessed anonymously, and participants provided informed consent before starting the survey. The recent study was conducted in accordance with the Declaration of Helsinki.

Consent for publication

Not applicable.

Competing interests

The authors declare no competing interests.

Author details

[1]Institute for General Practice and Interprofessional Care, University Hospital Tübingen, Osianderstr 5, 72076 Tübingen, Germany
[2]Institute of Applied Social Sciences, Technical University of Applied Sciences, Würzburg-Schweinfurt, Tiepolostraße 6, 97070 Würzburg, Germany
[3]Institute for Philosophy, University of Bremen, Enrique-Schmidt-Str 7, 28359 Bremen, Germany
[4]Institute of Occupational Medicine, Social Medicine and Health Services Research, University Hospital Tübingen, Wilhelmstr 27, 72074 Tübingen, Germany

Received: 29 June 2023 / Accepted: 16 January 2024
Published online: 23 January 2024

References

1. Wallace W, Chan C, Chidambaram S, Hanna L, Iqbal FM, Acharya A, et al. The diagnostic and triage accuracy of digital and online symptom checker tools: a systematic review. NPJ Digit Med. 2022;5(1):118.
2. GmbH AH. For better health 2023. Available from: https://ada.com/app/.
3. Pairon A, Philips H, Verhoeven V. A scoping review on the use and usefulness of online symptom checkers and triage systems: how to proceed? Front Med (Lausanne). 2022;9:1040926.
4. Vuillaume LA, Turpinier J, Cipolat L, Dumontier T, Peschanski N, Kieffer Y, et al. Exploratory study: evaluation of a symptom checker effectiveness for providing a diagnosis and evaluating the situation emergency compared to emergency physicians using simulated and standardized patients. PLoS ONE. 2023;18(2):e0277568.
5. Schmieding ML, Mörgeli R, Schmieding MAL, Feufel MA, Balzer F. Benchmarking triage capability of symptom checkers against that of medical laypersons: survey study. J Med Internet Res. 2021;23(3):e24475.
6. Aboueid S, Meyer S, Wallace JR, Mahajan S, Chaurasia A. Young adults' perspectives on the Use of Symptom checkers for self-triage and Self-Diagnosis: qualitative study. JMIR Public Health and Surveillance. 2021;7(1):e22637.
7. Kopka M, Scatturin L, Napierala H, Furstenau D, Feufel MA, Balzer F, et al. Characteristics of users and nonusers of Symptom checkers in Germany: cross-sectional survey study. J Med Internet Res. 2023;25:e46231.
8. Morse KE, Ostberg NP, Jones VG, Chan AS. Use characteristics and triage acuity of a digital symptom checker in a large integrated health system: population-based descriptive study. J Med Internet Res. 2020;22(11):e20549.
9. Carmona KA, Chittamuru D, Kravitz RL, Ramondt S, Ramirez AS. Health information seeking from an intelligent web-based symptom checker: cross-sectional questionnaire study. J Med Internet Res. 2022;24(8):e36322.
10. Hurrelmann K, Klinger J, Schaeffer D. Gesundheitskompetenz der bevölkerung in Deutschland im zeitvergleich der jahre 2014 und 2020. Gesundheitswesen [Internet]. 2022.
11. Meyer AND, Giardina TD, Spitzmueller C, Shahid U, Scott TMT, Singh H. Patient perspectives on the usefulness of an artificial intelligence-assisted symptom checker: cross-sectional survey study. J Med Internet Res. 2020;22(1):e14679.
12. Marco-Ruiz L, Bønes E, de la Asunción E, Gabarron E, Aviles-Solis JC, Lee E, et al. Combining multivariate statistics and the think-aloud protocol to assess human-computer interaction barriers in symptom checkers. J Biomed Inform. 2017;74:104–22.
13. Winn AN, Somai M, Fergestrom N, Crotty BH. Association of use of online symptom checkers with patients' plans for seeking care. JAMA Netw Open. 2019;2(12):e1918561.
14. Kopka M, Schmieding ML, Rieger T, Roesler E, Balzer F, Feufel MA. Determinants of laypersons' trust in medical decision aids: randomized controlled trial. JMIR Hum Factors. 2022;9(2):e35219.
15. Aboueid S, Meyer SB, Wallace J, Chaurasia A. Latent classes associated with the intention to use a symptom checker for self-triage. PLoS ONE. 2021;16(11):e0259547.
16. Starcevic V, Berle D, Arnáez S. Recent insights into Cyberchondria. Curr Psychiatry Rep. 2020;22(11).
17. Von Elm E, Altman DG, Egger M, Pocock SJ, Gøtzsche PC, Vandenbroucke JP, et al. The strengthening the reporting of observational studies in epidemiology (STROBE) statement: guidelines for reporting observational studies. Int J Surg. 2014;12(12):1495–9.

18. Tennant B, Stellefson M, Dodd V, Chaney B, Chaney D, Paige S, et al. eHealth literacy and web 2.0 health information seeking behaviors among baby boomers and older adults. J Med Internet Res. 2015;17(3):e70.
19. Jungmann SM, Brand S, Kolb J, Witthöft M. Do Dr. Google and health apps have (comparable) side effects? An experimental study. Clin Psychol Sci. 2020;8(2):306–17.
20. Balapour A, Reychav I, Sabherwal R, Azuri J. Mobile technology identity and self-efficacy: implications for the adoption of clinically supported mobile health apps. J Inf Manag. 2019;49:58–68.
21. Aboueid S, Liu RH, Desta BN, Chaurasia A, Ebrahim S. The use of artificially intelligent self-diagnosing digital platforms by the general public: scoping review. JMIR Med Inform. 2019;7(2):e13445.
22. Nießen D, Groskurth K, Rammstedt B, Lechner CM. General Life Satisfaction Short Scale (L-1). Zusammenstellung sozialwissenschaftlicher Items und Skalen (ZIS). 2020.
23. Soellner R, Huber S, Reder M. The concept of eHealth literacy and its measurement: German translation of the eHEALS. J Media Psychol. 2014;26(1):29–38.
24. Glöckner-Rist A, Barenbrügge J, Rist F. Deutsche Version des Whiteley Index (WI-d). Zusammenstellung sozialwissenschaftlicher Items und Skalen (ZIS). 2014.
25. Speckens AE, Spinhoven P, Sloekers PP, Bolk JH, van Hemert AM. A validation study of the Whitely Index, the illness attitude scales, and the Somatosensory amplification scale in general medical and general practice patients. J Psychosom Res. 1996;40(1):95–104.
26. Beierlein C, Kovaleva A, Kemper C, Rammstedt B, editors. Allgemeine Selbstwirksamkeit Kurzskala (ASKU). Zusammenstellung Sozialwissenschaftlicher Items Skalen (ZIS); 2014.
27. Wessel D, Attig C, Franke T. ATI-S - An Ultra-Short Scale for Assessing Affinity for Technology Interaction in User Studies. Proceedings of Mensch und Computer. 2019; Sep; Hamburg, Germany: Association for Computing Machinery; 2019. p. 147– 54.
28. Tibshirani R. Regression shrinkage and selection via the lasso. J R Stat Soc B (Methodol). 1996;58(1):267–88.
29. Ho D, Imai K, King G, Stuart E, Whitworth A. Package 'MatchIt'. Version[Google Scholar]. 2018.
30. Ho DE, Imai K, King G, Stuart EA. Matching as nonparametric preprocessing for reducing model dependence in parametric causal inference. Political Anal. 2007;15(3):199–236.
31. VanderWeele TJ. Principles of confounder selection. Eur J Epidemiol. 2019;34:211–9.
32. Stekhoven DJ, Buhlmann P. MissForest–non-parametric missing value imputation for mixed-type data. Bioinformatics. 2012;28(1):112–8.
33. James G, Witten D, Hastie T, Tibshirani R. An introduction to statistical learning: Springer; 2013.
34. Efron B. Estimating the error rate of a prediction rule: improvement on cross-validation. J Am Stat Assoc. 1983;78(382):316–31.
35. R Core Team. R: a language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing; 2016.
36. R Core Team. R: a language and environment for statistical computing. 4.0.5 ed. Vienna, Austria: R Foundation for Statistical Computing; 2021.
37. RStudio Team. RStudio: integrated development for R. PBC. Boston: RStudio; 2021.
38. Chen H, Cohen P, Chen S. How big is a big odds ratio? Interpreting the magnitudes of odds ratios in epidemiological studies. Commun Stat Simul Comput. 2010;39(4):860–4.
39. Vismara M, Caricasole V, Starcevic V, Cinosi E, Dell'Osso B, Martinotti G, et al. Is cyberchondria a new transdiagnostic digital compulsive syndrome? A systematic review of the evidence. Compr Psychiatry. 2020;99:152167.
40. Starcevic V, Berle D, Arnáez S. Recent insights into cyberchondria. Curr Psychiatry Rep. 2020;22(11):56.
41. Arsenakis S, Chatton A, Penzenstadler L, Billieux J, Berle D, Starcevic V, et al. Unveiling the relationships between cyberchondria and psychopathological symptoms. J Psychiatr Res. 2021;143:254–61.
42. Turner J, Knowles E, Simpson R, Sampson F, Dixon S, Long J, et al. Impact of NHS 111 online on the NHS 111 telephone service and urgent care system: a mixed-methods study. Health Serv Delivery Res. 2021;9(21):1–148.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

2.4. Publication 4

Publication 4, titled "Do I really not need to go to the doctor?" - Unveiling usage patterns and predicting usage of Symptom Checker Apps: an explorative longitudinal study," investigates the circumstances under which medical laypeople employ SCA and highlights the aspects users find most significant after using these apps. This publication is currently under review by the Journal of Medical Internet Research (2023 Impact Factor: 7.4).

Original Paper

Unveiling Usage Patterns and Explaining Usage of Symptom Checker Apps: Explorative Longitudinal Mixed Methods Study

Anna-Jasmin Wetzel[1], MSc; Christine Preiser[2], PhD; Regina Müller[3], PhD; Stefanie Joos[1], Prof Dr Med; Roland Koch[1], MD; Tanja Henking[4], Prof Dr; Hannah Haumann[1], MD

[1]Institute of General Practice and Interprofessional Care, University Hospital Tübingen, Tübingen, Germany

[2]Institute of Occupational and Social Medicine and Health Services Research, University Hospital Tübingen, Tübingen, Germany

[3]Institute for Philosophy, University of Bremen, Bremen, Germany

[4]Institute for Applied Social Sciences, Technical University of Würzburg-Schweinfurt, Würzburg, Germany

Corresponding Author:
Anna-Jasmin Wetzel, MSc
Institute of General Practice and Interprofessional Care
University Hospital Tübingen
Osianderstr. 5
Tübingen, 72074
Germany
Phone: 49 70712985231
Email: jasmin.wetzel@med.uni-tuebingen.de

Abstract

Background: Symptom checker apps (SCA) aim to enable individuals without medical training to classify perceived symptoms and receive guidance on appropriate actions, such as self-care or seeking professional medical attention. However, there is a lack of detailed understanding regarding the contexts in which individuals use SCA and their opinions on these tools.

Objective: This mixed methods study aims to explore the circumstances under which medical laypeople use SCA and to identify which aspects users find noteworthy after using SCA.

Methods: A total of 48 SCA users documented their medical symptoms, provided open-ended responses, and recorded their SCA use along with other variables over 6 weeks in a longitudinal study. Generalized linear mixed models with and those without regularization were applied to consider the hierarchical structure of the data, and the models' outcomes were evaluated for comparison. Qualitative data were analyzed through Kuckartz qualitative content analysis.

Results: Significant predictors of SCA use included the initial occurrence of symptoms, day of measurement (odds ratio [OR] 0.97), self-rated health (OR 0.80, $P<.001$), and the following International Classification in Primary Care-2–classified symptoms, that are general and unspecified (OR 3.33, $P<.001$), eye (OR 5.56, $P=.001$), cardiovascular (OR 8.33, $P<.001$), musculoskeletal (OR 5.26, $P<.001$), and skin (OR 4.76, $P<.001$). The day of measurement and self-rated health showed minor importance due to their small effect sizes. Qualitative analysis highlighted four main themes: (1) reasons for using SCA, (2) diverse affective responses, (3) a broad spectrum of behavioral reactions, and (4) unmet needs including a lack of personalization.

Conclusions: The emergence of new and unfamiliar symptoms was a strong determinant for SCA use. Specific International Classification in Primary Care–rated symptom clusters, particularly those related to cardiovascular, eye, skin, general, and unspecified symptoms, were also highly predictive of SCA use. The varied applications of SCA fit into the concept of health literacy as bricolage, where SCA is leveraged as flexible tools by patients based on individual and situational requirements, functioning alongside other health care resources.

(J Med Internet Res 2024;26:e55161) doi: 10.2196/55161

KEYWORDS

self-triage; eHealth; self-diagnosis; mHealth; mobile health; usage; patterns; predicts; prediction; symptoms checker; apps; applications; explorative longitudinal study; self care; self management; self-rated; mixed method; circumstances; General Linear Mixed Models; GLMM; qualitative data; content analysis; Kuckartz; survey; participants; users

XSL•FO
RenderX

Introduction

Symptom checker apps (SCA) are an emerging class of health technology tools that enable patients to input symptoms and biodata to receive a list of probable diagnoses and related triage suggestions [1]. Research on SCAs has largely focused on their accuracy, for instance, a study by Semigran et al [2] evaluated 23 SCAs using patient vignettes, revealing that the correct diagnosis appeared as the most probable in only 34% of cases. The appropriateness of the triage advice was found to be satisfactory in 57% of cases [2]. Notably, the accuracy of SCAs in nonemergency situations decreased as the urgency of the presented scenarios decreased [2]. A subsequent audit in 2022 indicated that there had been no significant improvement in triage performance over the past 5 years, with considerable variability in effectiveness among different SCAs [3]. Comparisons between SCAs and the decision-making of nonprofessionals showed no superiority in triage capability by SCAs [4]. Another study using simulated patients' cases showed that SCAs correctly identified the primary diagnosis in 30% of cases, significantly less than emergency physicians, who had an 80% success rate [5]. Despite a slightly better performance in emergency triage situations [4], physicians still vastly outperformed SCAs, with a 96% accuracy rate compared with 71% for SCAs [5]. Users of SCAs are typically characterized as younger females with a higher socioeconomic status [6] and a greater affinity for technology [7]. In addition, it has been observed that SCA users showed higher levels of health anxiety than nonusers [8], with SCAs being more widely recognized among these users [6]. Users with anxiety disorders, however, tend to find SCAs less beneficial [6], indicating that SCAs might not be advantageous for this user group. The literature suggests that such patient groups may not only fail to benefit from using SCAs but could also experience heightened anxiety as a result, implying that it might be better to avoid these applications [9].

The impact of SCAs on health care providers is also critical, as patients may use these tools as a substitute for or in addition to professional health care services [10]. The level of trust users place in SCAs and their previous medical knowledge significantly affects how SCA results are interpreted [11,12]. SCA might also lead to the inefficient use of health care resources [5], as users may seek emergency department care too soon or too frequently. This behavior, especially from those with conditions that are not urgent, can intensify the burden on health care systems by potentially increasing costs and reallocating resources away from individuals who require immediate medical attention [13,14]. This outcome contrasts with the promises that SCAs would reduce pressure on health care systems. Currently, it is unclear whether SCAs have an impact on health care systems and, if so, the nature of this impact.

The factors driving SCA usage remain underexplored. Research has shown that gender, age, health anxiety, and other psychosocial factors significantly influence SCA use [6]. User characteristics, including health anxiety, acceptance of technology, and education level, also affect how SCAs are used [15,16]. Despite this knowledge, there is a notable gap in research regarding the situational factors that prompt SCA usage. Gaining a deeper insight into these situational factors that lead to SCA use is essential for comprehending the broader implications of SCA use on both individuals and health care providers. Consequently, this study uses a mixed methods approach to explore how medical laypersons use SCAs, aiming to understand the patterns of use and the users' perceptions following their experience with a symptom checker.

Methods

Overview

This study was conducted using a mixed methods study approach, incorporating both qualitative and quantitative components within a prospective longitudinal cohort study design. The primary outcome of this study was whether one specific SCA was used during a day or not by Ada [17]. Ada [17] was chosen for the study due to its stage of development and its professed broad usage. Moreover, the application is ISO (International Organization for Standardization)-certified and possesses the CE mark. The use of one specific SCA was analyzed to ensure standardized interaction with one SCA only. The STROBE (The Strengthening the Reporting of Observational Studies in Epidemiology) guidelines were applied [18]. This study was conducted as a part of the joint project CHECK.APP [19].

In order to answer the research questions (predictors and usage patterns of SCA, aspects noteworthy after usage) of this mixed methods study, different longitudinal mixed models for SCA use were explored and compared.

Ethical Considerations

Ethical approval was granted by the ethics committee of the medical faculty of the University of Tuebingen (464/202BO). Each participant was compensated 300€ (US $358) for their participation in the study. In accordance with the ethical approval, all participants provided written consent before taking part in the user diaries. A written and verbal explanation was provided, and participation in the study was voluntary.

Measurements

User diaries were filled out daily by the study participants over a period of 6 weeks. SCA use was binary coded (yes or no). Further, the user diary comprised a survey that contained different questions about self-rated mood, self-rated health, self-rated stress level, and occurrence of physical symptoms throughout the day. Self-rated health and mood were both assessed with a visual analog 5-point scale. Subjectively rated stress was assessed with a single-choice item and 3 answer possibilities. If participants had symptoms, they stated what kind of symptoms they had, whether they took medication, and if they used a specific SCA [17]. If participants used the SCA on a day this was classified as a use-case, if the SCA was not used on a day this was classified as a nonuse case. Participants were asked if the symptoms appeared for the first time. Those who reported symptoms were asked to describe them in open-ended responses. After the data acquisition, reported symptoms were classified according to the International Classification in Primary Care 2 (ICPC) by a doctor from the study team, using dummy coding (present not present). Age

XSL•FO
RenderX

was assessed in years, and gender was assessed as an open-ended response resulting in the 2 categories “male” and “female.” The variable “day of study” was retrieved from the dates in the user diaries.

Further questions considering the procedure after the app use were recorded as well. If the SCA was used, open-ended responses in the user diary asked for the experience with the SCA.

Data Collection

The participants of the study were recruited from December 2020 to March 2021, with data collection taking place from April 2021 to September 2021. Due to the exploratory nature of the study, the sample size was determined based on estimated effect sizes and the feasibility of data collection. The requirement of the ethics committee for the study was not to initiate the use of SCA, and therefore only to include participants in the study who were already using SCA before the study. The participants were recruited from a previously conducted survey of the CHECK.APP project, specifically those who reported using a specific SCA [17]. The recruitment process for the survey involved using mailing lists from the University of Tuebingen, a social media campaign by the University Hospital Tuebingen, and social media channels from a statutory health insurance (Allgemeine Ortskrankenkasse Baden-Wuerttemberg). Participants were then contacted through the mail and asked whether they wanted to participate in a further study part.

Prior to the study’s start, web-based workshops were conducted to train participants on how to use the user diary and to enable them to ask questions to prevent common mistakes. Participants were directed to use the SCA as they normally would in their everyday activities. They chose when to start their 6-week period and submitted their diaries after finishing it. About 3 months after the workshop, they were sent reminder emails.

Inclusion and Exclusion Criteria

The inclusion criteria for the study included having at least a B2 level of proficiency in the German language, as classified by the Common European Framework of Reference for Languages, an age older than 18 years, and being an Ada user. The exclusion criterion was possessing a medical degree.

Data Analysis

Quantitative Data

To analyze and process the quantitative data, R statistics (version 4.1.3; R Core Team) [20] and RStudio (version 1.4) [20] were used. The aim of the present mixed methods study is to investigate the context in which SCAs are used by medical laypeople. To reach this aim, different longitudinal mixed models for SCA use are explored and compared.

Generalized Linear Mixed Model

To consider the hierarchical data generalized linear mixed models (GLMMs) with regularization, using a least absolute shrinkage and selection operator (LASSO) penalization, and without regularization were fit to the data and compared. The R package glmmPen [21] was used to conduct the LASSO penalized GLMM, and unpenalized model parameters were derived using the R package lme4 [22]. Further explanation can be found in Multimedia Appendix 1 and Textbox 1.

The model assessment was realized by using the Bayesian information criterion (BIC) scores of the models.

If model 3 or 4 would be identified as best performing they were fit with a GLMM without regularization to derive parameter estimators, CI, and SE.

Textbox 1. Mixed models.

- Model 0: Intercept-only model, with a random intercept for participants
- Model 1: Generalized linear mixed models based on expert knowledge containing only fixed effects for the variables:
  - Day of measurement
  - First-time appearance of symptoms
  - Self-rated health
- Model 2: as model 1 with added random effects for the variables “self-rated health,” “first time appearance of symptoms”
- Model 3: Regularized generalized linear mixed models with least absolute shrinkage and selection operator penalty containing all predictors of Table 1, with independent variance-covariance matrix
- Model 4: as model 3 with unstructured variance-covariance matrix

XSL•FO
RenderX

Table 1. Coefficients, CI, and *P* values for Model 0-3.

Coefficient	Model 0[a]		Model 1[b]		Model 2[c]		Model 3[d]	
	OR (95% CI)	*P* values	OR (95% CI)	*P* values	OR (95% CI)	*P* values	OR (95% CI)	*P* values
Intercept	0.14 (0.12-0.17)	<.001	3.49 (1.76-6.92)	<.001	5.18 (2.41-11.13)	<.001	359.87 (55.38-2338.31)	<.001
Day of measurement	__[e]	—	0.97 (0.96-0.98)	<.001	0.97 (0.96-0.98)	<.001	0.97 (0.96-0.99)	<.001
Self-rated health	—	—	0.44 (0.37-0.52)	<.001	0.40 (0.32-0.49)	<.001	0.80 (0.64-1.01)	<.001
Initial occurrence of Symptoms (yes)	—	—	9.28 (6.06- 14.22)	<.001	13.95 (6.20-31.37)	<.001	5.97 (3.67-9.72)	<.001
General and unspecified (present)	—	—	—	—	—	—	3.33 (1.92-5.56)	<.001
Eye (present)	—	—	—	—	—	—	5.56 (2.08-12.29)	<.001
Cardiovascular (present)	—	—	—	—	—	—	8.33 (2.70-25.00)	<.001
Musculoskeletal (present)	—	—	—	—	—	—	5.26 (3.44-8.33)	<.001
Skin (present)	—	—	—	—	—	—	4.76 (2.04-11.11)	<.001

[a]σ^2=3.29; τ_{00}=0.18$_{ID}$; Bayesian information criterion=1568.02; intraclass correlation coefficient=0.05; N=48 (2016 observations).

[b]σ^2=3.29; τ_{00}=0.26$_{ID}$; Bayesian information criterion=1306.17; intraclass correlation coefficient=0.07; N=48 (2016 observations).

[c]σ^2=3.29; τ_{00}=0.40$_{Intercept}$; τ_{11}=0.03 (participant-rated health) and 3.26 (first-time appearance of symptoms); ρ_{01}=–0.65 and –0.06; Bayesian information criterion=1328.20; intraclass correlation coefficient=0.13; N=48 (2016 observations).

[d]σ^2=3.29; τ_{00}=1.14$_{Intercept}$; τ_{11}=1.53 (general and unspecified); ρ_{01}=–0.89$_{ID}$; Bayesian information criterion=1134.82; intraclass correlation coefficient=0.13; N=48 (2016 observations).

[e]Not applicable.

Missing Data

An Overview of missing data on the outcome and predictor variables can be found in Multimedia Appendix 2. All variables had less than 5% missing data. Missing data on the primary outcome and predictors were imputed using the R package missForest [23], that enables the imputation of missing in mixed data (categorical and continuous), based on a random forest approach [23]. The out-of-bag error (proportion of falsely classified) was used to quantify the test error.

Qualitative Data Analysis

The written comments from the open-ended questions were analyzed with qualitative content analysis [24]. Main categories and subcategories were built inductively after initial familiarization with the data [25]. All transcripts were analyzed by JW using MaxQDA [26]. For quality control, 40% of the codings were independently coded by CP. Afterward, the few conflicting codings were discussed by both coders in data sessions.

Results

Descriptive Data

Of the 48 participants, 31 (71%) were female. The mean age of the participants was 27 (SD 9.1, range 19-64) years. The primary outcome variable showed 264 use cases and 1752 nonuse cases. Imputation of missings resulted in an out-of-bag error of 0.063. Symptoms of the chapter blood had low cell frequencies, the variable “Blood” was therefore excluded as a predictor in the following data analysis for the primary outcome. Tables with frequencies stratified by the outcome of the predictor variables can be found in Multimedia Appendices 3 and 4.

Quantitative Results: Primary Outcome and Model Comparison

All assumptions of GLMM were met. Binned residual Plots for Models 1-3 can be found in Multimedia Appendix 5. The Null model revealed an intraclass correlation of 0.05 and indicated no substantial evidence of clustering between the participants. The BIC of the Intercept-only model was BIC=1568.02. Model 1 derived a BIC of 1305.00. The BIC achieved by Model 2 was higher (BIC=1326.76) compared with the BIC of Model 1. The BIC of Model 3 (BIC=1165.60) revealed the best model performance with the lowest value. Therefore, this model was chosen over Model 4 (BIC=1180.12) and was fit using a GLMM without regularization to derive parameter estimators, CI and SE.

LASSO regularization of Model 3 identified 8 fixed effect predictors, which are the day of measurement, self-rated health, the first-time appearance of symptoms, and 5 ICPC classified symptom clusters (general and unspecified, eye, cardiovascular, skin, and musculoskeletal) as meaningful. All predictors except self-rated health revealed a significant *P*<.001.

An overview of parameter estimations, 95% CI, and the BIC can be found in Table 1.

XSL•FO
RenderX

Qualitative Results

The qualitative content analysis revealed four different main categories: (1) reasons for using SCA (n=79), (2) affective response and evaluation considering SCA usage (n=122), (3) Behavioral response considering SCA usage (n=66), and (4) unmet requirements (n=33).

As users made several entries into the diaries, topics varied between users in general and between the days the entry was made.

Reasons for Using Symptom Checker Apps

Reasons for SCA use emerged as a key topic. Participants identified several reasons for using SCA, including gaining information about medical conditions, symptoms, and treatment options. Participants also used symptom checkers to classify their symptoms and to understand the potential underlying causes of their health issues. The latter is not only to their own symptoms but also to the symptoms of close relatives and friends. Another stated reason for using a SCA was to determine whether a visit to a healthcare professional was necessary or not. Participants also reported using symptom checkers to cross-reference their own assumptions about their health concerns against the information provided by the SCA.

> *For validation: It's probably "really just" migraine & my head isn't about to burst!*

In addition, SCA was used for verifying medical diagnoses and gaining a better understanding of one's health concerns, which was seen as helpful in order to make informed decisions about one's health and well-being. Moreover, participants used already known diseases or conditions to check if the SCA was able to detect them correctly and see how the SCA performed.

> *After the COVID-19 vaccination, I experienced side effects such as fever, headache, and dizziness. I knew that these symptoms were related to the vaccination, but I still wanted to know if the SCA could attribute these symptoms to the vaccination, which it did.*

Affective Response and Evaluation Considering the Usage of Symptom Checker Apps

Participants' affective responses and evaluations following their use of symptom checkers varied widely. While some participants expressed in the respective entries positive or satisfied feelings and evaluations, others reported negative or dissatisfied emotions and evaluations, and some were ambivalent.

> *Woke up with neck pain, according to [the SCA] it could be a tension. [The SCA] recommended heat and stretching, which helped. App usage easy, some questions were alarming.*

Positive evaluations were associated with various facets of SCA usage. These include feelings of satisfaction and contentment, perceiving the information offered by the SCA as helpful and suitable, experiencing a heightened sense of security, achieving a better grasp of one's symptoms, enriching the user's knowledge, and finding the SCA's responses to be logical, realistic, or plausible. Affective responses to SCA use encompass feeling calm, relieved, or more relaxed, feeling sensitized, and motivated to take action. Furthermore, users reported an increased sense of agency and empowerment.

> *I found it very interesting to read about this topic. I notice that the [SCA] is also helpful and useful to me, even for minor physical complaints.*

Participants' negative and dissatisfied emotions and evaluations were associated with several aspects of SCA use. Some participants reported having no trust in the symptom checker, finding the information provided unrealistic or unlikely, and feeling that the SCA worsened their situation if they were already experiencing health anxiety. Others found the information provided to be partially or entirely inaccurate or unhelpful, and some found the recommendations exaggerated or irritating. The negative emotions experienced by participants included disappointment, shock, disillusionment, or feeling unsure, anxious, scared, alarmed, or worried, and some expressed surprise or disappointment in the results provided by the SCA.

> *The [SCA] gives me three possible causes: tonsillitis, peritonsillar abscess, and retropharyngeal abscess. For all three, it recommends visiting the emergency room. I think it's exaggerated to go to the emergency room for a sore throat since I don't have a fever.*

Behavioral Response Considering the Usage of Symptom Checker Apps

Participants reported various behaviors following their use of SCA. In some cases, participants searched the internet for more detailed information, while others preferred to wait and observe their symptoms before deciding whether to seek medical care. Self-treatment strategies were frequently mentioned, such as using over-the-counter medicine, increasing fluid intake, relaxing, or using home remedies.

> *I did not follow the advice of the [SCA] app to seek medical advice as I want to treat it myself with the ointment for a few days first.*

In some cases, participants contacted the health care system and either considered making a doctor's appointment, immediately made a doctor's appointment, or decided to address their symptoms during their next scheduled appointment. However, some participants explicitly stated that they did not plan to discuss their SCA use with a doctor. A few participants made it clear that they chose not to follow the guidance provided by the SCA.

> *[The SCA] advises going to the emergency room, no possibility to see a doctor, was very worried throughout the day, wanted to go directly to the emergency room after the day, but symptoms have subsided a bit, so waited until the next day.*

Unmet Requirements

Participants in this study identified a number of unmet requirements when using the SCA. Some participants expressed the need for a more complex symptom input system that allows for a more nuanced description of symptoms, as well as the ability to differentiate between different qualities and quantities of symptoms.

XSL•FO
RenderX

> *My symptoms were both diverse and unspecific - therefore [SCA] couldn't really make any concrete suggestions - as good as the help texts are, sometimes the questions asked are difficult to answer - I miss [the SCA] not knowing me better from the beginning, so I have fed her with pre-existing conditions, I am curious to see how well Ada learns from that. It's a shame that this cannot yet be taken into account in the suggestions.*

Others reported that their symptoms were not recognized by the SCA and that there was no category for their specific complaint. Participants also expressed concern over the fact that symptoms and diagnoses (eg, bronchitis or cough) can be entered into the SCA. Participants expressed that the weighting of symptoms should be based on the leading symptom entered. Some participants found the questions difficult to answer and reported redundant questions. Participants requested a more differentiated classification of symptoms, as well as the ability to recognize different causes of multiple entered symptoms. Participants evaluated the level of personalization in the symptom checker as insufficient, as preexisting medical conditions, lifestyle, family history, and medication use were not considered.

> *Since I am young, exercise a lot, drink little, do not smoke, and generally live very healthily, I rarely had significant complaints. When something did occur, it was mostly self-inflicted (hangover feeling, sun, sports injuries). I would have wished that not only symptoms but also more lifestyle decisions had been addressed in the assessment. Most often, these explained my symptoms very accurately.*

Some participants also requested the recognition of medication side effects and interactions, as well as consideration of lifestyle factors such as alcohol consumption and sleep deprivation. In addition, participants suggested that the SCA provide recommendations for home remedies and medication use. Participants also requested more educational material that illustrates the relationship between symptoms and specific advice. Finally, participants were concerned about the unclear source of the information provided by the SCA.

Discussion

Principal Results

This mixed methods study aimed to understand the patterns of SCA usage and to describe its use among medical laypersons. The new appearance of unfamiliar symptoms was found to be a strong predictor for SCA usage. In addition, specific ICPC symptom clusters such as “cardio-vascular,” “eye,” “skin,” “musculoskeletal,” and “general and unspecified symptoms” were also associated with a higher probability of SCA use. The qualitative data analysis identified 4 main themes reflecting the aspects that participants found noteworthy after using an SCA. The main themes comprised (1) reasons for using SCA, (2) diverse affective responses, (3) a broad spectrum of behavioral reactions, and (4) unmet needs including a lack of personalization.

Comparison to Previous Work

To evaluate SCA use in a broader context it is worthwhile to reflect on why patients go to see a physician. Braunacker-Mayer and Avery [27] described that the decision to seek help from a physician depends on complex aspects of social, psychological, cultural, and biomedical factors. She described patients as “non-passive users of health care services”; who do not visit doctors without consideration. Further, she claims that patients evaluate their own health condition and then take steps to manage symptoms or treat illnesses. Thus, the decision-making process for seeking medical assistance may not align with clinicians’ perspectives, as it is influenced by a broader range of factors beyond biomedical considerations. With regard to the limited understanding of the SCA usage context, it seems beneficial to explore how the extended version of the Andersen model, a well-established framework in health care usage research, can provide insights into this aspect. The Andersen model [28] incorporates three dimensions to analyze the factors influencing health care usage. These dimensions include predisposing characteristics (such as age and gender), enabling resources (like access to health care facilities and income), and need factors (such as the prevalence of physical illnesses and perceived health status). A further study suggested expanding the model to include psychosocial factors, such as loneliness, and personality traits, like neuroticism [29]. In the following sections, we will examine and classify the identified predictors in the context of SCA usage and health care usage, with qualitative data being integrated subsequently. In the following paragraphs, we use the Andersen model on health care usage to contextualize our results.

Predictor “Initial Occurrence of Symptoms”

In this study, we found that the predictor “initial occurrence of symptoms” had a large effect on SCA use. In line with the Andersen model for health care usage [28] this predictor might represent a need factor. Furthermore, it can be considered as both situational and personal as the “initial occurrence of symptoms” relies on specific symptoms presented by an individual and includes the fact that this individual had experienced similar symptoms before or not. The perception of novel or familiar symptoms relies on prior experiences and the personal capacity to distinguish between different symptoms as well as the overall evaluation of symptoms. It is remarkable that other data have revealed anxiety as a predictor for SCA use in a cross-sectional setting [6]. Health anxiety might have an influence on the appraisal of “initially occurred symptoms” as relevant or irrelevant. Therefore, psychosocial factors or personality traits not only may have an influence on face-to-face health care usage but also on SCA use.

Through qualitative data analysis, we gained deeper insights into the underlying needs related to the predictor “initial occurrence of symptoms.” The category “purpose of use” in the qualitative results shed light on users’ motivations for using SCA. Users reported using these applications to gather information about medical conditions, symptoms, and treatment options. They also wanted to assess their own symptoms or those of close individuals to better understand potential underlying health issues. Another reason for using SCA was to

XSL•FO
RenderX

determine whether a visit to a healthcare professional was necessary. These findings indicate that SCA contributes to the bricolage of health literacy. This perspective defines health literacy as a multidimensional concept, emerging from situational, dynamic, and social practices [30]. Health care bricolage involves the creative assembly and usage of a broad spectrum of resources, including a variety of knowledges, ideas, materials, and networks, to manage specific health concerns [31]. SCA seems to adapt and serve distinct functions, using diverse sources and types of health knowledge, along with the interplay between social connections and situational factors. They can be used by patients as a flexible resource tailored to their personal and situational needs, parallel to other resources.

Predictor "Self-Rated Health"

Compared to previous findings regarding positive associations between health care usage and self-rated health [32], self-rated health was identified as a predictor with a minor negative effect on SCA use. Regarding the association between need factors in the Andersen model [28] of health care usage, substantial evidence demonstrated a positive link between these factors [33]. In particular, self-rated health showed a strong association with health care usage. This finding is highly plausible since an increase in health care needs often indicates illness symptoms that prompt individuals to consult a physician, as supported by numerous cross-sectional studies. This relationship may not apply in the context of SCA use as they are a lower threshold offer and users' needs are therefore less connected to illness symptoms and may be strongly connected to curiosity or insecurities raised by the initial appearance of symptoms and health anxiety and personality traits as mentioned above. In another cross-sectional study, self-rated health was examined in the context of inclination to use SCA, and a positive association was found [6]. Interestingly, no association was found between finding SCA useful and self-rated health [6]. Nevertheless, the recent findings concerning the relationship between SCA and self-rated health remain inconclusive, emphasizing the need for further research considering the inclination and the actual usage of SCA to delve into this aspect more comprehensively.

Predictor "Day of Measurement"

A further predictor with an uncertain role and ambiguous impact was the "day of measurement," which indicated that the timing within the study significantly influenced SCA usage. However, the effect size of this predictor was too small to meaningfully influence the outcome. It should be noted that the large number of measurements may have contributed to the significance of the predictor.

Predictors Classified by International Classification in Primary Care

Other predictors examined were symptoms entered into the SCA categorized according to Symptoms of the categories "cardiovascular," "general and unspecified," "musculoskeletal," "eye," and "skin" were selected as relevant predictors, and all revealed a decent effect size and significant ORs. All named symptom clusters are reported as frequent reasons for emergency consultations in different studies [34-36].

The OR of the predictor "cardiovascular" symptoms was found to be the highest in our study, indicating that the usage of SCA is associated with potentially threatening symptoms. From a qualitative study in our project, it emerged that users would not use SCA in acute emergencies, such as a traumatic brain injury after a traffic accident. However, the findings suggest that users do use SCA for potentially severe and, therefore, anxiety-inducing symptoms. This could be explained by a form of internal triage by medical laypersons. In the case of an obviously nondeferrable emergency, SCA is not used. However, when uncertainties exist, and symptoms are anxiety-inducing but not immediately classifiable as time-sensitive and life-threatening, SCA is used. However, a study from 2022 [35] revealed that SCA does not exhibit a notable performance in such critical scenarios. This should be considered a high-risk situation for patients who may rely on SCA recommendations that are potentially wrong. "General and nonspecific" symptoms often point to a range of possible conditions, posing challenges in their attribution. Consequently, they could serve as meaningful predictors for SCA usage, as users seek additional information to determine the severity and urgency of such symptoms. This need for further clarification underscores the importance of SCA in assisting users to classify and understand their symptoms accurately.

The appearance of symptoms of the cluster "musculoskeletal," "eye," and "skin" frequently leads to a high level of distress as they are restricting in everyday life and therefore might lead to SCA usage. "Musculoskeletal" symptoms lead to movement restrictions and are therefore frequently perceived as constraining in everyday life. A study reports them as the main symptoms of attending an ED [36]. Symptoms concerning the eyes are often perceived as fear-inducing and irritating [37]. Moreover, symptoms concerning the skin are commonly visible and could therefore be stigmatizing [37]. They are also reported as symptoms with a low subjectively perceived treatment urgency that, however, lead patients to seek emergency consultations [34].

The identified ICPC symptom clusters are in accordance with the 10 most common general practitioner consultation reasons reported by the CONTENT project [38]. This may indicate 2 aspects: First, the need for low-threshold medical advice for these symptom clusters; Second SCAs are used for symptoms that may lead to contact with the medical system, and SCAs presumably have an impact on health care professionals and further resources.

Strengths and Limitations

The study investigated the use of a specific SCA over a 42-day period and achieved a 0% dropout rate due to compensation, regular reminders, and user workshops. As a result, the data had less than 5% missing values. Since all participants used the same SCA to ensure comparability, our results may not be generalizable to all available SCAs. Due to the low frequencies of certain ICPC-2 chapters, some predictors were excluded, which may have resulted in some relevant but rarely occurring predictors being left out of the data. Another crucial aspect to consider with the ICPC-2 is its inability to reveal symptom severity or urgency. As a result, the recent findings do not take

XSL•FO
RenderX

these factors into account. It should be noted that LASSO penalization in variable selection comes with certain general limitations. LASSO tends to favor sparse models over complex ones, resulting in models that are more interpretable [39]. However, in cases where there is a high degree of multicollinearity, this preference for sparsity can lead to instability in predictor selection [39]. While we anticipate some multicollinearity among predictors, particularly due to overlaps between ICPC-coded symptoms, overall, multicollinearity is not expected to be a significant issue. LASSO is generally effective at handling moderate multicollinearity [39]. In addition, variable selection in LASSO is influenced by the choice of the shrinking term and penalty [39], which represents an inherent limitation of the method. Nonetheless, compared with other variable selection techniques such as forward or backward selection, LASSO proves to be better suited for the presented data, particularly when considering moderate multicollinearity and the volume of data available.

Moreover, the sample consisted predominantly of females and younger adults, raising uncertainty about its representativeness of the entire user demographic or whether it reflects a sampling bias Nonetheless, this user sample aligns with other findings regarding the sociodemographic characteristics associated with SCA usage [6]. Consequently, additional studies encompassing broader user groups and diverse use cases are imperative to extrapolate the findings more universally.

Conclusion

In this study, SCAs were more frequently used in situations characterized by greater uncertainty for example in situations when symptoms appeared for the first time. This suggests that the concept of intolerance to uncertainty may influence SCA use. Further research is essential to explore the relationships between intolerance of uncertainty, health anxiety, and SCA use as this aspect was not addressed in this study. In addition, the symptom clusters identified in our study had a large overlap with symptoms reported in emergency department visits considering recent study results, indicating that SCA usage could potentially affect health care usage. Further research is needed to ascertain whether such effects exist and how they impact health care usage. Moreover, SCAs are not only used intentionally and purposefully; they are also used playfully, for reasons beyond personal use and as this study showed to validate the user's or doctor’s hypotheses. Furthermore, they are used to verify the accuracy of the software itself, especially when a diagnosis is already known. This reveals that SCAs contribute to a broader bricolage of health literacy, serving as one of many tools integrated into a wider context of sources of information regarding one’s health. Given the current low usage rate of SCAs in Germany, this effect may though be minor at present in the German health care system. Nevertheless, various stakeholders in the German health care system have shown interest in SCAs. Currently, it remains uncertain how this interest will affect SCA usage rates in the future.

Acknowledgments

We thank Johannes A. Vey and Malte Klemmt for their advice. Further, we would like to thank the participants of the cross-sectional survey. We used the generative artificial intelligence tool ChatGPT (OpenAI) to revise the language and style of the manuscript, which were further reviewed and revised by the study group.

This study was funded by the Federal Ministry of Education and Research (BMBF), which is a cabinet-level ministry of Germany (grant 01GP1907A). The funder played no role in the study design, data collection, data analysis, and interpretation, or manuscript writing.

Data Availability

The datasets generated during and/or analyzed during this study are available from the corresponding author on reasonable request.

Authors' Contributions

AJW, HH, RK, and SJ designed the study and the study materials. JW collected the data and performed the analyses. HH verified the quantitative data and contributed to the interpretation of the results. CP verified the qualitative data and contributed to the interpretation of the results. AJW wrote the first draft of the manuscript, with critical input from RM, SJ, CP, TH, and HH. All authors reviewed and approved the final manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Execution for the generalized linear mixed models.
[DOCX File , 15 KB-Multimedia Appendix 1]

Multimedia Appendix 2

Overview of the missings.
[DOCX File , 19 KB-Multimedia Appendix 2]

XSL•FO
RenderX

Multimedia Appendix 3

Overview of the symptom checker app use or nonuse cases considering potential contributing variables.
[DOCX File , 12 KB-Multimedia Appendix 3]

Multimedia Appendix 4

International Classification in Primary Care 2–coded symptoms stratified for symptom checker app use or nonuse.
[DOCX File , 38 KB-Multimedia Appendix 4]

Multimedia Appendix 5

Binned residual plot for Model 1-3.
[DOCX File , 253 KB-Multimedia Appendix 5]

References

1. Wallace W, Chan C, Chidambaram S, Hanna L, Iqbal FM, Acharya A, et al. The diagnostic and triage accuracy of digital and online symptom checker tools: a systematic review. NPJ Digit Med. 2022;5(1):118. [FREE Full text] [doi: 10.1038/s41746-022-00667-w] [Medline: 35977992]
2. Semigran HL, Linder JA, Gidengil C, Mehrotra A. Evaluation of symptom checkers for self diagnosis and triage: audit study. BMJ. 2015;351:h3480. [FREE Full text] [doi: 10.1136/bmj.h3480] [Medline: 26157077]
3. Schmieding ML, Kopka M, Schmidt K, Schulz-Niethammer S, Balzer F, Feufel MA. Triage accuracy of symptom checker apps: 5-year follow-up evaluation. J Med Internet Res. 2022;24(5):e31810. [FREE Full text] [doi: 10.2196/31810] [Medline: 35536633]
4. Schmieding ML, Mörgeli R, Schmieding MAL, Feufel MA, Balzer F. Benchmarking triage capability of symptom checkers against that of medical laypersons: survey study. J Med Internet Res. 2021;23(3):e24475. [FREE Full text] [doi: 10.2196/24475] [Medline: 33688845]
5. Abensur Vuillaume L, Turpinier J, Cipolat L, Arnaud-Dépil-Duval, Dumontier T, Peschanski N, et al. Exploratory study: evaluation of a symptom checker effectiveness for providing a diagnosis and evaluating the situation emergency compared to emergency physicians using simulated and standardized patients. PLoS One. 2023;18(2):e0277568. [FREE Full text] [doi: 10.1371/journal.pone.0277568] [Medline: 36827277]
6. Kopka M, Scatturin L, Napierala H, Fürstenau D, Feufel MA, Balzer F, et al. Characteristics of users and nonusers of symptom checkers in Germany: cross-sectional survey study. J Med Internet Res. 2023;25:e46231. [FREE Full text] [doi: 10.2196/46231] [Medline: 37338970]
7. Aboueid S, Meyer SB, Wallace J, Chaurasia A. Latent classes associated with the intention to use a symptom checker for self-triage. PLoS One. 2021;16(11):e0259547. [FREE Full text] [doi: 10.1371/journal.pone.0259547] [Medline: 34731217]
8. Wetzel A, Klemmt M, Müller R, Rieger MA, Joos S, Koch R. Only the anxious ones? Identifying characteristics of symptom checker app users: a cross-sectional survey. BMC Med Inform Decis Mak. 2024;24(1):21. [FREE Full text] [doi: 10.1186/s12911-024-02430-5] [Medline: 38262993]
9. Starcevic V, Berle D, Arnáez S. Recent insights into cyberchondria. Curr Psychiatry Rep. 2020;22(11):56. [FREE Full text] [doi: 10.1007/s11920-020-01179-8] [Medline: 32852626]
10. Radionova N, Ög E, Wetzel A, Rieger MA, Preiser C. Impacts of symptom checkers for laypersons' self-diagnosis on physicians in primary care: scoping review. J Med Internet Res. 2023;25:e39219. [FREE Full text] [doi: 10.2196/39219] [Medline: 37247214]
11. Woodcock C, Mittelstadt B, Busbridge D, Blank G. The impact of explanations on layperson trust in artificial intelligence-driven symptom checker apps: experimental study. J Med Internet Res. 2021;23(11):e29386. [FREE Full text] [doi: 10.2196/29386] [Medline: 34730544]
12. Kopka M, Schmieding ML, Rieger T, Roesler E, Balzer F, Feufel MA. Determinants of laypersons' trust in medical decision Aids: randomized controlled trial. JMIR Hum Factors. 2022;9(2):e35219. [FREE Full text] [doi: 10.2196/35219] [Medline: 35503248]
13. Pairon A, Philips H, Verhoeven V. A scoping review on the use and usefulness of online symptom checkers and triage systems: how to proceed? Front Med (Lausanne). 2022;9:1040926. [FREE Full text] [doi: 10.3389/fmed.2022.1040926] [Medline: 36687416]
14. Turner J, Knowles E, Simpson R, Sampson F, Dixon S, Long J, et al. Impact of NHS 111 online on the NHS 111 telephone service and urgent care system: a mixed-methods study. Health Services and Delivery Research. 2021;9(21):1-148. [Medline: 34780129]
15. Aboueid S, Meyer S, Wallace JR, Mahajan S, Chaurasia A. Young adults' perspectives on the use of symptom checkers for self-triage and self-diagnosis: qualitative study. JMIR Public Health Surveill. 2021;7(1):e22637. [FREE Full text] [doi: 10.2196/22637] [Medline: 33404515]

XSL•FO
RenderX

16. Pairon A, Philips H, Verhoeven V. A scoping review on the use and usefulness of online symptom checkers and triage systems: how to proceed? Front Med (Lausanne). 2022;9:1040926. [FREE Full text] [doi: 10.3389/fmed.2022.1040926] [Medline: 36687416]
17. Ada. Über uns. 2020. URL: https://ada.com/de/about/ [accessed 2020-12-10]
18. von Elm E, Altman DG, Egger M, Pocock SJ, Gøtzsche PC, Vandenbroucke JP, et al. STROBE Initiative. The strengthening the reporting of observational studies in epidemiology (STROBE) statement: guidelines for reporting observational studies. Int J Surg. 2014;12(12):1495-1499. [FREE Full text] [doi: 10.1016/j.ijsu.2014.07.013] [Medline: 25046131]
19. Wetzel AJ, Koch R, Preiser C, Müller R, Klemmt M, Ranisch R, et al. Ethical, legal, and social implications of symptom checker apps in primary health care (CHECK.APP): protocol for an interdisciplinary mixed methods study. JMIR Res Protoc. 2022;11(5):e34026. [FREE Full text] [doi: 10.2196/34026] [Medline: 35576570]
20. Team RC. R: A language and environment for statistical computing. 4.0.5 ed. Vienna, Austria. R Foundation for Statistical Computing; 2021.
21. Heiling HM, Rashid NU, Li Q, Ibrahim JG. glmmPen: high dimensional penalized generalized linear mixed models. R J. 2023;15(4):106-128. [FREE Full text] [doi: 10.32614/rj-2023-086] [Medline: 38818017]
22. Bates D, Mächler M, Bolker B, Walker S. Fitting linear mixed-effects models using lme4. J. Stat. Soft. 2015;67(1):1-48.
23. Stekhoven DJ, Bühlmann P. MissForest--non-parametric missing value imputation for mixed-type data. Bioinformatics. 2012;28(1):112-118. [doi: 10.1093/bioinformatics/btr597] [Medline: 22039212]
24. Kuckartz U. Qualitative text analysis. Los Angeles, LA. SAGE Publications; 2013.
25. Kuckartz U. Qualitative Inhaltsanalyse. Methoden, Praxis, Computerunterstützung (Grundlagentexte Methoden). Germany. Beltz Juventa; 2012.
26. MAXQDA 2020. Berlin, Germany. VERBI Software; 2019.
27. Braunack-Mayer A, Avery JC. Before the consultation: why people do (or do not) go to the doctor. Br J Gen Pract. 2009;59(564):478-479. [doi: 10.3399/bjgp09x453495]
28. Andersen RM. Revisiting the behavioral model and access to medical care: does it matter? J Health Soc Behav. 1995;36(1):1-10. [Medline: 7738325]
29. Hajek A, König H-H. Beyond symptoms: why do patients see the doctor? BJGP Open. 2020;4(2):bjgpopen20X101088. [FREE Full text] [doi: 10.3399/bjgpopen20X101088] [Medline: 32430301]
30. Samerski S. Health literacy as a social practice: Social and empirical dimensions of knowledge on health and healthcare. Soc Sci Med. 2019;226:1-8. [doi: 10.1016/j.socscimed.2019.02.024] [Medline: 30831555]
31. Phillimore J, Bradby H, Knecht M, Padilla B, Pemberton S. Bricolage as conceptual tool for understanding access to healthcare in superdiverse populations. Soc Theory Health. 2018;17(2):231-252. [doi: 10.1057/s41285-018-0075-4]
32. Xu F, Johnston JM. Self-rated health and health service utilization: a systematic review. Int J Epidemiol. 2015;44(suppl_1):i180.
33. Hajek A, Kretzler B, König HH. Determinants of healthcare use based on the andersen model: a systematic review of longitudinal studies. Healthcare (Basel). 2021;9(10):1354. [FREE Full text] [doi: 10.3390/healthcare9101354] [Medline: 34683034]
34. Scherer M, Lühmann D, Kazek A, Hansen H, Schäfer I. Patienten in Notfallambulanzen: Querschnittstudie zur subjektiv empfundenen Behandlungsdringlichkeit und zu den Motiven, die Notfallambulanzen von Krankenhäusern aufzusuchen. Dtsch Arztebl Int. 2017;114(39):645-652. [FREE Full text] [doi: 10.3238/arztebl.2017.0645] [Medline: 29034865]
35. Fraser HSF, Cohan G, Koehler C, Anderson J, Lawrence A, Pateña J, et al. Evaluation of diagnostic and triage accuracy and usability of a symptom checker in an emergency department: observational study. JMIR Mhealth Uhealth. 2022;10(9):e38364. [FREE Full text] [doi: 10.2196/38364] [Medline: 36121688]
36. Unwin M, Kinsman L, Rigby S. Why are we waiting? Patients' perspectives for accessing emergency department services with non-urgent complaints. Int Emerg Nurs. 2016;29:3-8. [FREE Full text] [doi: 10.1016/j.ienj.2016.09.003] [Medline: 27776980]
37. Okosa MC, Uwakwe R, Apakama AI, Onwuegbuna AA, Uzozie CC, Amobi MBC, et al. Psychological effects of eye diseases: a tertiary eye center study. J Psychiatry Psychiatr Disord. 2021;05(04):128-139.
38. Laux G, Kühlein T, Gutscher A, Szecsenyi J. Versorgungsforschung in der Hausarztpraxis. München. Urban und Vogel; 2010.
39. Hastie T, Tibshirani R, Friedman JH, Friedman JH. The elements of statistical learning: data mining, inference, and prediction. Berlin, Heidelberg. Springer; 2009.

Abbreviations

BIC: Bayesian information criterion
GLMM: Generalized linear mixed model
ICPC: International Classification in Primary Care 2
ISO: International Organization for Standardization
LASSO: least absolute shrinkage and selection operator

XSL•FO
RenderX

SCA: symptom checker apps
STROBE: Strengthening the Reporting of Observational Studies in Epidemiology

Edited by A Mavragani; submitted 07.12.23; peer-reviewed by K Savolainen, H Napierala, M Kopka; comments to author 30.01.24; revised version received 13.03.24; accepted 14.10.24; published 09.12.24

Please cite as:
Wetzel A-J, Preiser C, Müller R, Joos S, Koch R, Henking T, Haumann H
Unveiling Usage Patterns and Explaining Usage of Symptom Checker Apps: Explorative Longitudinal Mixed Methods Study
J Med Internet Res 2024;26:e55161
URL: https://www.jmir.org/2024/1/e55161
doi: 10.2196/55161
PMID:

©Anna-Jasmin Wetzel, Christine Preiser, Regina Müller, Stefanie Joos, Roland Koch, Tanja Henking, Hannah Haumann. Originally published in the Journal of Medical Internet Research (https://www.jmir.org), 09.12.2024. 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 work, first published in the Journal of Medical Internet Research (ISSN 1438-8871), is properly cited. The complete bibliographic information, a link to the original publication on https://www.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

3. Discussion

The following paragraphs will answer the research questions from chapter 1.5. Subsequently in chapter 3.4 an overarching chapter triangulating all results from this thesis with concepts of health anxiety will be discussed.

3.1. Overview and Impact of SCA Usage

What is the extent of SCA usage in Germany and which applications are most prevalent?

In examining the usage of SCA in Germany, our research reveals a moderate utilization rate, with 7.5% of respondents reporting usage[35], closely aligning with the 6.5% reported by other German research groups[16]. Contrary to expectations, the most prevalent SCA was the NetDoktor Symptom Checker, followed by Ada. Contrary to expectations, Ada's significant presence may be largely attributed to its very active advertising strategy, which initially led to predictions that it would be the most prevalent software.

How do SCA integrate within the HISB framework?

Regarding the integration of SCA within the HISB framework, SCA appear to play a minor role compared to other information sources. Predominantly, individuals consult doctors or turn to the internet, with SCA being the least utilized across all age demographics[35]. One explanation is the general lack of knowledge about the existence of SCA in Germany[35]. Notably, SCA users tend to seek information more frequently online compared to non-users[35]. Surprisingly, the age group of 51-55 showed the highest SCA usage[35], deviating from the expected predominance in younger demographics as suggested by Meyer et al[28]. This deviation might come from younger individuals preferring direct internet searches over using SCA, viewing them as competing resources.

Is there a correlation between SCA usage and the frequency of consultations with GP, and patient satisfaction with these medical professionals?

Findings from this dissertation indicate no significant correlation between SCA usage and the frequency of consultations with GP or patient satisfaction levels with these medical professionals[35]. Since the recent results were not evaluated

in an experimental setting, causal conclusions cannot be drawn. Based on the current findings, there is no indication that SCA users turn to SCA due to dissatisfaction with their GP.

What are the factors contributing to the non-use of SCA?

Among the non-users, 25% were aware of SCA, while a majority of 67% participants had never heard of it[35]. More than half (54%) of the non-users expressed no interest in trying SCA[35]. Several factors contribute to SCA non-use. One reason is, there is a widespread lack of knowledge regarding the existence of SCA[35]. Further doubts about their trustworthiness were reported. Additionally, SCA are perceived as an inadequate initial point for healthcare by non-users, and there are concerns that they cannot handle the complexity of medical diagnostic processes[35]. Further barriers include the lack of necessary technical competencies to use SCA effectively, as well as skepticism about the substantial benefits derived from their use[35].

3.2. Predictors of SCA Usage

Which user characteristics predict the use of SCA? How do these characteristics influence the adoption and sustained use of SCA?

Health anxiety emerged as the strongest predictor of SCA usage[36]. When the Whiteley Index was categorized with a cut-off score greater than 5, 52% of SCA users met the criteria for hypochondria, compared to only 8% of non-users[36]. This level of health anxiety may impair the ability of patients to act appropriately on recommendations and symptom classifications provided by SCA, potentially leading to negative outcomes. The use of SCA for self-diagnosis among individuals prone to cyberchondria was advised against in a 2020 study, and the importance of managing expectations when accessing health information online was highlighted[37]. Furthermore, it was indicated by another study that individuals with high hypochondria scores often felt worse after using online symptom checkers, whereas reassurance was experienced by those with lower scores[38]. These observations, coupled with our research, suggest that SCA are more likely to be used by patients with health anxiety, yet less likely to benefit them, emphasizing

the complex relationship between user characteristics and the effectiveness of SCA.

Additionally, lower self-efficacy scores were initially identified as a positive predictor in the main analysis, although this effect was not replicated in subsequent sensitivity analyses[36]. In a 2023 study, patients with higher self-efficacy reported finding SCA more useful, although this did not extend to knowledge about them or the intention to use them[16]. The role of self-efficacy in the context of SCA remains unclear and must be further investigated to draw definitive conclusions.

3.3. Situational Factors in SCA Adoption

What situational factors influence the adoption of SCA among medical laypeople?

Significant predictors of SCA use included the initial occurrence of symptoms, the day of measurement, self-rated health, and symptoms classified under the ICPC-2 (International Classification of Primary Care) categories: general and unspecified, eye, cardiovascular, musculoskeletal, and skin[39]. However, the day of measurement and self-rated health were considered of minor importance due to their small effect sizes[39].

The predictor initial occurrence of symptoms can be considered both situational and personal, as it depends on the specific symptoms presented by an individual and includes whether similar symptoms had been experienced previously. The perception of symptoms as novel or familiar is influenced by prior experiences, the personal capacity to distinguish between different symptoms, and the overall evaluation of symptoms. Remarkably, other studies have shown that anxiety is a predictor for SCA use in a cross-sectional setting[16], and health anxiety may influence the appraisal of initially occurring symptoms as relevant or irrelevant. Thus, it is suggested that psychosocial factors or personality traits influence not only face-to-face healthcare utilization but also SCA use.

These identified ICPC symptom clusters align with the ten most common GP consultation reasons reported by the CONTENT project[40]. This alignment indicates a need for low-threshold medical advice for these symptom clusters and suggests

that SCA are utilized for symptoms that may lead to contact with the medical system, presumably impacting healthcare professionals and further resources.

What aspects of SCA use are most significant to users after they have utilized these tools?

Four main themes that participants found noteworthy after using an SCA were revealed. These themes include (1) purpose of SCA use, (2) diverse affective responses, (3) a broad spectrum of behavioral reactions, and (4) unmet needs including a lack of personalization[39]. The category purpose of use from the qualitative results sheds light on users' motivations for utilizing SCA, highlighting their use of these applications to gather information about medical conditions, symptoms, and treatment options, to assess their own or others' symptoms to better understand potential underlying health issues, and to determine the necessity of visiting a healthcare professional[39].

3.4. Health Anxiety in the Context of SCA Use

In the research presented in this dissertation, a consistent theme of health anxiety associated with the use of SCA emerged. The upcoming chapter seeks to triangulate the relationship between the scientific findings obtained throughout this dissertation and concepts related to health anxiety.

Definitions of key concepts

The subsequent paragraph will outline the concepts related to health anxiety relevant to this thesis: health anxiety, hypochondria, intolerance of uncertainty (IU) and cyberchondria.

Health anxiety is defined as inappropriate or excessive fear and worry about a perceived threat to one's health[41]. It can be regarded on a continuum, ranging from no health concerns to severe, pathological anxiety about health[42, 43]. Excessive and maladaptive health anxiety is a central aspect of hypochondria and is frequently observed in various other disorders, including somatoform, anxiety, and mood disorders[42]. Unlike the spectrum-based understanding of health anxiety, the diagnosis of hypochondria is treated as categorical[44]. Hypochondria is known to lead to higher healthcare utilization, including frequent doctor visits and

increased sick leave[45, 46], as individuals believe they may have a severe disease. Although the terms 'health anxiety' and 'hypochondria' are partially interchangeably used in literature[46], their definitions are debated, with scientists and clinicians treating them as distinct conditions or as health anxiety as a continuous concept with hypochondria as an endpoint[42]. The DSM-IV (Diagnostic and Statistical Manual of Mental Disorders) classified hypochondria as a somatoform disorder characterized by the misinterpretation of somatic sensations, a categorization that faced criticism for its emphasis on somatic symptoms over illness anxiety [47]. Consequently, in the DSM-V, hypochondria was replaced and can now be categorized into two different diagnoses depending on the presence and intensity of somatic symptoms[48]. If patients exhibit pronounced somatic symptoms that trigger health anxieties and concerns, a diagnosis of Somatic Symptom Disorder is made[48]. Illness Anxiety Disorder in DSM-V is diagnosed when somatic symptoms are absent or only mildly present[48]. Thus, the DSM-V distinguishes pathological health anxieties based on the presence and severity of somatic symptoms[49]. Unlike the ICD-10 (International Classification of Diseases) and DSM-IV diagnosis of Hypochondria, DSM-V includes not just the fear of having a serious illness but also the fear of acquiring one [49]. The division of hypochondria into two diagnoses has also faced criticism. The separation based on the intensity of somatic symptoms is seen as artificial and unsupported by scientific evidence[49]. In the general population, the 1-year prevalence of Hypochondria (according to DSM-III, DSM-IV, and ICD-10) ranges from 0 to 4.5%[50], and the point prevalence in Germany is 0.4% [51]. Consequently, up to 4.5% of the general population suffers from Hypochondria within a year.

IU is defined as “the dispositional incapacity of an individual to endure the adverse reaction triggered by the perceived absence of key and sufficient information and supported by the associated perception of uncertainty”[52]. The concept of IU is known as a vulnerability factor for the development of anxiety disorders, including health anxiety and associated concepts as Hypochondria and Cyberchondria[53]. Further, results from 2023 revealed that illness-anxious individuals report higher levels of anxiety sensitivity and IU and interpret ambiguous health-related situations more catastrophically[54].

Cyberchondria is characterized by the repetitive searching for health-related information online, leads to increased health anxiety levels[55]. Although not currently recognized as a distinct disorder[56,57], the debate over whether it should be considered a separate disorder is ongoing[58]. It is linked to increased healthcare utilization and functional impairments. Cyberchondria involves various elements: the time-consuming and repetitive nature of online health searches, negative emotional states and physiological reactions triggered by these searches, internal conflicts regarding the trustworthiness of online information versus professional medical advice, the unwanted nature of these searches, and the pursuit of reassurance[59]. Multiple studies have identified moderate to strong correlations between cyberchondria, health anxiety[56,57], and hypochondria[60]. It has been observed that managing feelings of uncertainty plays a critical role in cyberchondria. Specifically, an inability to tolerate uncertainty is thought to contribute to cyberchondria and the behavior of seeking reassurance[61]. Uncertainty can actually increase following online health searches, particularly when the information found is ambiguous, inconsistent, or contradictory. Such uncertainty prompts further searches in an effort to reduce doubts and achieve a sense of "closure." The link between cyberchondria and difficulty in tolerating uncertainty is underscored by findings that show intolerance of uncertainty to be a significant predictor of health anxiety triggered by searching for health information on the Internet [62].

In this dissertation a continuous concept of health anxiety was assumed with an endpoint of hypochondria. The Whiteley Index (WI) was used to assess health anxiety, and it was considered in a continuous as well as a dichotomized version. Health anxiety was used as a concept to refer to the continuous WI values and Hypochondria to the dichotomized WI. It was chosen to assess health anxiety resp. Hypochondria instead of Cyberchondria, Somatic Symptom Disorder or Illness Anxiety Disorder as more overarching themes were of interest in the exploratory study setting. Literature of Cyberchondria was also considered, due to the overlapping of the concepts and thematic proximity resp. the overlapping of SCA use and online research. IU was a concept that emerged during data interpretation and seems to play an important role not only in the context of health anxiety, hypochondria and cyberchondria but also in the context of SCA use.

Contextualizing health anxiety within the findings of this thesis

As part of publication three hypochondria was assessed using the WI, which established it as a reliable predictor of characteristics among SCA users[36]. Moreover, using a WI cut-off score of five[63], 52% of SCA users were classified under hypochondria[36], a significantly higher proportion compared to 8% among non-SCA users [6] and 4.5% among the German population[50].

In addition, the qualitative analysis of publication three showed that SCA users expressed concerns about the potential impact of SCA use, e.g. increasing their health anxiety[36]. A finding in the category affective response of the qualitative data analysis was the observation that SCA usage may exacerbate conditions for those already suffering from health anxiety[39]. One participant stated:

> *"I am a bit of a hypochondriac and I fear that if I used such devices, I would think I was ill even more often than I do now." (240)*

In publication four the initial occurrence of symptoms was pinpointed as a key factor leading to SCA use[39]. Other predictors, including cardiovascular, general, or unspecified symptoms, depend largely on the patient's interpretation as they can be interpreted ambiguously and can be potentially life-threatening[39]. The process of symptom evaluation as novel or ambiguous is critical, especially for patients with heightened health anxiety, where symptom appraisal and IU play significant roles[54].

SCA use and health anxiety - Comparison to prior work

In another study within the CHECK.APP research project, user interviews uncovered a sub-category labeled Apprehension[64]. This highlights health anxiety as a motivator for using SCA, thus underscoring the role of health anxiety in the adoption of SCA[64]. Another identified sub-category in the same study was Concerns, which encompassed mental distress symptoms such as nervousness, impatience, and anxiety experienced during the use of SCA[64]. For instance, one participant noted that pre-existing anxieties could be exacerbated by SCA usage. However, the existing body of literature on the relationship between SCA and health anxiety remains limited. One study found an increased utilization of SCA among individuals with panic and anxiety disorders. Additionally, this group

reported experiencing SCA as less beneficial more often than not [16]. The findings suggest that while individuals with panic and anxiety disorders may be more inclined to use SCA, they frequently do not find these apps helpful[16]. Other researchers have recommended that individuals suffering from Cyberchondria should be discouraged from using SCA, arguing that such usage could lead to an undue focus on diagnosis and self-diagnosis, practices that should be avoided when suffering from Cyberchondria[11]. In contrast, a study sponsored by a software development found that SCA users experienced less health-related anxiety after using a specific SCA[24].

In the course of this dissertation, evidence suggested that IU also may influence the use of SCA in several ways: Users of SCA were more likely to have higher levels of health anxiety, which is associated with IU; and SCA were more often utilized in situations marked by uncertainty, such as when symptoms first occurred and participants could not revert to strategies they had previously used or when symptoms were ambiguous, as seen in the ICPC-symptom category General and unspecified. These findings imply that IU may impact SCA use in these contexts.

To manage IU, over-engagement behaviors are one strategy individuals use to increase certainty by engaging in overpreparation and extensive information seeking[65]. While over-engagement is not dysfunctional by nature, its rigid use can increase IU[65]. Individuals who favor over-engagement as a method to reduce uncertainty may use SCA as a form of over-engagement. Given that higher levels of IU are associated with anxiety disorders and related concepts, this reliance on SCA can potentially heighten anxiety and healthcare utilization, presenting further challenges for patients and healthcare providers.

Another important factor is that individuals with health anxiety exhibit a lower tolerance for general ambiguous situations and tend to interpret health-related ambiguities less safely compared to controls[54]. Considering that individuals with higher levels of health anxiety are more likely to use SCA and considering SCA frequently present multiple possible diagnoses and various recommendations of action, this combination appears particularly dysfunctional. Additionally, it has

been observed that individuals with health anxiety are more sensitive to physical symptoms[54], suggesting they might have a greater number of situations in which they find using SCA relevant compared to individuals with lower levels of health anxiety. Figure 1 illustrates the integration of SCA into the framework of health anxiety and IU. It visualizes potential negative outcomes through SCA use for individuals experiencing health anxiety. Potential benefits, such as reduced health anxiety through reassurance from an SCA output, are possible though not illustrated in Figure 1. Health anxiety may decrease if the SCA suggests only non-critical diagnoses and provides consistent triage advice for all suggested diagnoses with a low urgency level. However, recent studies have shown that individuals with high hypochondria scores often feel worse after using online symptom checkers, while those with lower scores are more likely to experience reassurance[38]. This difference may be influenced by IU, which affects how individuals with health anxiety perceive and appraise health information, making it more challenging for them to feel reassured[54]. Additionally, SCA tend to perform less reliably in self-care scenarios, resulting in less frequent low-level triage advice in these cases[12].

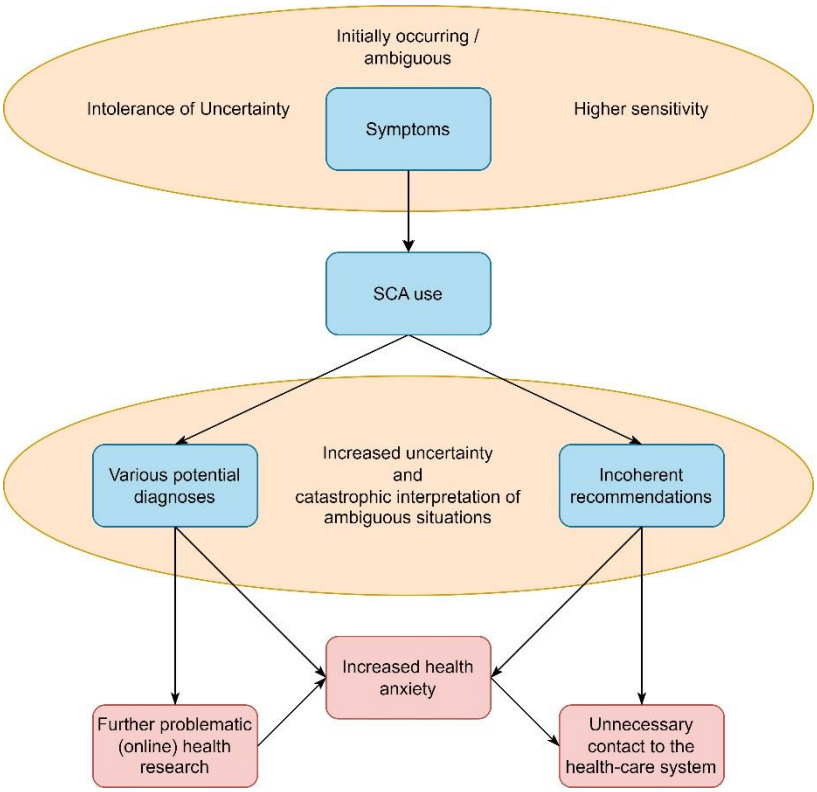


Figure 1: Potential negative outcomes of SCA use for individuals with health anxiety. Blue – SCA use, Orange – Inner processes, Red - Negative consequences.

Relationship between health-care utilization, health-anxiety and SCA usage

The literature on the impact of SCA on healthcare utilization remains sparse. There are concerns that SCA might contribute to the inefficient use of healthcare resources[8], as individuals could prematurely or excessively seek emergency care. This behavior, notably from those with non-urgent conditions, may exacerbate healthcare system burdens by potentially driving up costs and distracting resources from those in need of urgent care[15,26]. Such outcomes are in contrast to the anticipated benefits of SCA, which were expected to relieve pressures on healthcare systems. The potential for SCA to increase healthcare system burdens could be raised from various factors. A primary concern is the insufficient performance of SCA delivering inaccurate recommendations[9,66,67], that may lead to inappropriate patient actions. Additionally, the ambiguity generated by SCA offering varying diagnoses and recommendations could contribute to this issue.

This is particularly problematic for SCA users with higher levels of health anxiety, who are already more inclined to utilize a broad spectrum of healthcare services, including primary and specialist care. Inconsistent advice from SCA may heighten uncertainty and anxiety, also resulting in further unnecessary healthcare utilization[68]. The precise impact of SCA on healthcare systems, as well as the nature of this impact, remains uncertain but requires further investigation, particularly in the context of a health-anxious user group.

In Germany, SCA currently appear to have a minor role, as indicated by a low usage rate of 7-8% and a general lack of awareness about SCA[16]. Given the modest interest among the German population and the increasing availability of SCA through various healthcare stakeholders (e.g., health insurance companies[20], hospital groups[19], National Association of Statutory Health Insurance Physicians[18]), the future influence of SCA on healthcare utilization and the system at remains unclear. The evolution of SCA's popularity and impact over the coming years is an open question, with potential implications for healthcare delivery and policy.

Recommendations for SCA in the context of health anxious users

Users with higher health anxiety are more inclined to use SCA, yet they tend to benefit less from their use. This can be the case because SCA can exacerbate health anxiety rather than provide reassurance[38], especially when the apps offer inconclusive recommendations and diagnostic suggestions. To mitigate increased health anxiety among vulnerable SCA users, it might be beneficial for SCA to incorporate specific features designed to protect and guide these users. This could include warning notices upon download or initial use, highlighting that while SCA can offer general health information, they also may increase health anxiety. Additionally, prompts or alerts for users exhibiting signs of excessive use on SCA could encourage consultation with healthcare professionals, providing a safeguard against the potential for increased anxiety or misdiagnosis. Implementing such prompts requires a nuanced understanding of the user base, particularly those most vulnerable to the negative impacts of SCA. Collaboration with healthcare professionals in developing these guidelines can ensure that users are directed towards reliable, professional care when needed. Research is

necessary to develop and evaluate the efficacy and practicality of this potential mechanisms.

3.5. Strengths and Limitations

Strengths

This dissertation provides a comprehensive analysis of SCA, examining the subject from multiple viewpoints. This includes an overarching perspective, insights into user characteristics, and situational factors, offering a multifaceted understanding of SCA. Additionally, the study of various symptom checker software enhances the generalizability of the findings. The findings of this thesis provide valuable insights for caregivers, users, software developers, and policymakers to consider when evaluating the practical use of SCA and potential enhancements.

Further, all publications were written by an interdisciplinary team, enabling the integration of diverse perspectives and methodologies. The primary goal of the dissertation was to give a preliminary overview of the field, marking a significant first step in exploration.

Methodologically, the presented results adopted a mixed-methods approach, combining quantitative and qualitative data analyses to uncover diverse aspects of the topic. This methodological choice allows for a broader investigation of SCA usage and impact. Additionally, the research incorporated diverse study designs, utilizing both cross-sectional and longitudinal data to enable comprehensive insights and enhance the validity of the results. For the quantitative approach, both frequentist principles and machine learning methods were employed to achieve more comprehensive and robust results.

Limitations

All publications used an exploratory approach, focusing on research questions rather than testing specific hypotheses. This exploratory nature is both a strength and a limitation, as it opens new areas of investigation but also means that the research does not provide clear evidence to all questions posed.

Given the relevance of health anxiety, a more nuanced conceptual assessment of hypochondria or cyberchondria, alongside further research into the specific

consequences of SCA usage, would be advantageous. Additionally, exploring potential mechanisms to protect vulnerable groups in experimental settings could be beneficial.

Sampling presented challenges, particularly in reaching a diverse and representative group of participants. Recruitment was conducted through mailing lists, social media channels of the University Hospital Tübingen, the university itself, and the social media platforms of AOK, a health insurance company. While these efforts were methodologically acknowledged and adjustments were made to account for potential biases, the sampling process's limitations may affect the generalizability of the findings.

Large Language Models such as ChatGPT, which can serve as SCA, were not investigated in this dissertation. ChatGPT, which serves as one prominent example, was launched at the end of 2022, after the data collection period for this study had concluded. Consequently, this dissertation did not explore it as a potential example of an SCA. Further research is necessary to gain insights into the use of ChatGPT and other Large Language Models as SCA.

Further, the applicability of the results across different countries may be limited, particularly in regions where healthcare systems, primary care provision, and cultural and social conditions vary significantly from Germany.

3.6. Conclusions

SCA currently play a minor role in Germany, and their future impact is yet to be uncertain. With a variety of stakeholders behind these apps, transparency regarding the motivations and biases that may influence the apps' recommendations is essential. This is particularly important as the array of interests behind these apps could affect the quality and objectivity of the information provided. The current performance of SCA remains inadequate, potentially resulting in unsafe triage recommendations or contributing to unnecessary healthcare burdens. This could further strain an already resource-scarce primary care sector. Moreover, there is a risk of SCA causing overdiagnosis, thereby posing additional risks to patients and contradicting the initial promises of these apps. Addressing these challenges

necessitates enhancing SCA performance and better integrating them with primary care and emergency services.

Users with high health anxiety are more likely to turn to SCA, yet they often derive less benefit. Rather than providing reassurance, the use of SCA can intensify health anxiety, especially because SCA often offer inconclusive recommendations and diagnostic suggestions.

In conclusion, the current evidence indicates that SCA are not fulfilling their advertised potential to improve healthcare for users. The German population remains divided between skepticism and interest in SCA. Furthermore, it is unclear whether SCA are cost-effective or if they ultimately increase healthcare expenditures.

4. Summary

Background

This dissertation examines Symptom Checker Apps (SCA) in Germany from various perspectives. These apps represent a new type of digital tool designed to help users categorize their medical complaints and navigate the healthcare system. However, significant disparities in the performance of these apps' triage and diagnostic recommendations often lead to inadequate results. This raise concerns that users might either seek unnecessary medical care or, conversely, fail to seek necessary treatment. This dissertation explores three main themes using a mixed-methods study design: 1. Overview and impact of SCA usage, 2. User characteristics as predictors of SCA usage, and 3. Situational factors influencing SCA adoption.

Methodology

Manuscript 1 outlines the research design of the research project CHECK.APP through a study protocol. Manuscript 2 analyzes the distribution of SCA, potential correlations to primary care provision, and reasons against the use of SCA, using descriptive and univariate analyses as well as a thematic analysis. Manuscript 3 employs a logistic regression to identify predictive variables for SCA users. In Manuscript 4, the usage patterns of SCA users over six weeks are examined, using mixed- model comparisons and a qualitative content analysis according to Kuckartz, which categorizes user comments.

Results

Manuscript 2 found that 7.5% of the German population uses SCA, with 50% of non-users showing no interest in trying them. Surprisingly, the highest number of SCA users was in the age group of 51-55 years. No correlations were found between primary care provision and SCA. The most well-known SCA was identified as the NetDoktor Symptom Checker. Manuscript 3 identified health anxiety as a reliable predictor for the use of SCA. Manuscript 4 determined the initial occurrence of symptoms and the ICPC-2 clusters: general and unspecified symptoms,

eyes, cardiovascular, musculoskeletal, and skin as relevant predictors with sufficient effect size.

Discussion

Currently, SCA play a minor role in Germany, and the interest of the general population is divided. It remains uncertain what impact an increasing use of SCA could have on health care. When using SCA, concepts such as health anxiety and Intolerance of Uncertainty seem to play a role. SCA are used by users with higher health anxiety in situations of higher uncertainty, e.g., when new or ambiguous symptoms occur. At the same time, these users seem to benefit less from SCA.

5. German summary

Hintergrund

Die vorliegende Dissertation untersucht Symptom-Checker-Apps (SCA) in Deutschland aus verschiedenen Blickwinkeln. Diese Apps stellen eine neue Art digitaler Werkzeuge dar, die darauf abzielen, Nutzer:innen bei der Einordnung medizinischer Beschwerden zu unterstützen und den adäquaten Zugang zum Gesundheitssystem zu finden. Momentan gibt es jedoch erhebliche Leistungsunterschiede bei den Triage- und Diagnoseempfehlungen der SCA, die oft zu unzureichenden Ergebnissen führen. Dies wirft Bedenken auf, dass Nutzer:innen möglicherweise unnötige medizinische Versorgung in Anspruch nehmen oder umgekehrt notwendige medizinische Hilfe nicht in Anspruch nehmen. Die Dissertation untersucht drei Hauptthemen mittels eines gemischt-methodischen Ansatzes: 1. Überblick und Auswirkungen der Nutzung von SCA, 2. Nutzer:innen-Merkmale als Prädiktoren für die Nutzung von SCA und 3. Situative Faktoren, die die Nutzung von SCA beeinflussen.

Methodik

Manuskript 1 legt durch ein Studienprotokoll das Forschungsdesign des Forschungsprojekts CHECK.APP dar. In Manuskript 2 wird die Verbreitung von SCA, potenzielle Zusammenhänge zur hausärztlichen Versorgung und Gründe gegen die Nutzung von SCA analysiert, mittels deskriptiver und univariater Analysen sowie einer thematischen Analyse. Manuskript 3 verwendet eine logistische Regression, um prädiktive Variablen für SCA Nutzer:innen zu identifizieren. In Manuskript 4 werden über sechs Wochen die Nutzungsmuster von SCA-Nutzer:innen untersucht, unter Anwendung von Mehrebenenmodell-Vergleichen und einer qualitativen Inhaltsanalyse nach Kuckartz.

Ergebnisse

Manuskript 2 ergab, dass 7,5 % der deutschen Bevölkerung SCA nutzen, wobei 50 % der Nicht-Nutzer:innen kein Interesse an der Nutzung angaben. Überraschenderweise gab es die meisten SCA Nutzer:innen in der Altersgruppe zwischen 51-55 Jahren. Es fanden sich keine Zusammenhänge zur hausärztlichen

Versorgung und SCA. Als bekannteste SCA wurde der NetDoktor Symptom Checker identifiziert. Manuskript 3 identifizierte Gesundheitsangst als zuverlässigen Prädiktor für die Nutzung von SCA. In Manuskript 4 wurden das erstmalige Auftreten von Symptomen sowie die ICPC-2 Cluster: Allgemein und unspezifische Symptome, Augen, Kardiovaskulär, Muskuloskelettal und Haut als relevante Prädiktoren mit aussagekräftiger Effektstärke identifiziert.

Diskussion

SCA spielen derzeit in Deutschland eine untergeordnete Rolle, und das Interesse der Allgemeinbevölkerung ist geteilt. Es bleibt ungewiss, welche Auswirkungen eine zunehmende Nutzung von SCA auf die Primärversorgung haben würde. Bei der Nutzung von SCA scheinen die Konzepte Gesundheitsangst und Intolerance of Uncertainty eine Rolle zu spielen. SCA werden vermehrt von Nutzer:innen mit höherer Gesundheitsangst in Situationen mit höherer Unsicherheit z.B. beim Auftreten neuer oder mehrdeutigen Symptome verwendet. Gleichzeitig scheinen genau diese Nutzer:innen weniger von SCA zu profitieren.

6. References

1. Statista. Anzahl der im Apple App Store verfügbaren mHealth Apps vom 1. Quartal 2015 bis zum 3. Quartal 2020. Published January 01, 2024; Accessed May 20, 2024. https://de.statista.com/statistik/daten/studie/1191205/umfrage/anzahl-der-bei-apple-verfuegbaren-mhealth-apps/
2. Google. Googgle Play Store Search for “symptom checker”. Accessed July 10, 2020. https://play.google.com/store/search?q=symptom%20checker&c=apps&hl=de
3. Lau T, Osterloh F. Ärztemangel: Schlechte Aussichten. *Dtsch Arztebl Int*. 2022;119(22-23): A-1008 / B-842. https://www.aerzteblatt.de/archiv/225635/Aerztemangel-Schlechte-Aussichten.
4. Link E, Baumann E. Use of health information on the internet: personal and motivational influencing factors. *Bundesgesundheitsblatt-Gesundheitsforschung-Gesundheitsschutz*. 2020;63:681-689. doi:10.1007/s00103-020-03144-5
5. Wallace W, Chan C, Chidambaram S, et al. The diagnostic and triage accuracy of digital and online symptom checker tools: a systematic review. *NPJ Digit Med*. 2022;5(1):118. doi:10.1038/s41746-022-00667-w
6. Ada Health. For better health. Accessed December 10, 2023. https://ada.com/app/
7. Semigran HL, Linder JA, Gidengil C, Mehrotra A. Evaluation of symptom checkers for self diagnosis and triage: audit study. *BMJ*. 2015;351:h3480. doi:10.1136/bmj.h3480
8. Vuillaume LA, Turpinier J, Cipolat L, et al. Exploratory study: Evaluation of a symptom checker effectiveness for providing a diagnosis and evaluating the situation emergency compared to emergency physicians using simulated and standardized patients. *Plos One*. 2023;18(2). doi:10.1371/journal.pone.0277568
9. Schmieding ML, Kopka M, Schmidt K, Schulz-Niethammer S, Balzer F, Feufel MA. Triage Accuracy of Symptom Checker Apps: 5-Year Follow-up Evaluation. *J Med Internet Res*.2022;24(5):e31810. doi:10.2196/31810
10. Hill MG, Sim M, Mills B. The quality of diagnosis and triage advice provided by free online symptom checkers and apps in Australia. *Med J Aust*. 2020;212(11):514-519. doi:10.5694/mja2.50600
11. Hill MG. *Appraisal of free online symptom checkers and applications for self-diagnosis and triage: an Australian evaluation*. Master thesis. Edith Cowan University Perth; 2020.
12. Schmieding ML, Mörgeli R, Schmieding MAL, Feufel MA, Balzer F. Benchmarking Triage Capability of Symptom Checkers Against That of Medical Laypersons: Survey Study. *J Med Internet Res*. 2021;23(3):e24475. doi:10.2196/24475
13. Kopka M, Feufel MA, Berner ES, Schmieding ML. How suitable are clinical vignettes for the evaluation of symptom checker apps? A test

theoretical perspective. *Digit Health.* 2023;9:20552076231194929. doi:10.1177/20552076231194929
14. Pairon A, Philips H, Verhoeven V. A scoping review on the use and usefulness of online symptom checkers and triage systems: How to proceed? *Front Med*. 2022;9:1040926. doi:10.3389/fmed.2022.1040926
15. Turner J, Knowles E, Simpson R, et al. Impact of NHS 111 Online on the NHS 111 telephone service and urgent care system: a mixed-methods study. *Health Serv and Deliv Res*. 2021;9(21):1-148. doi: 10.3310/hsdr09210
16. Kopka M, Scatturin L, Napierala H, et al. Characteristics of Users and Nonusers of Symptom Checkers in Germany: Cross-Sectional Survey Study. *J Med Internet Res*. 2023;25:e46231. doi:10.2196/46231
17. TK legt Kooperation mit Ada Health auf Eis. Updated December 05, 2019. Accessed May 15, 2020. https://www.aerztezeitung.de/Wirtschaft/TK-legt-Kooperation-mit-Ada-Health-auf-Eis-404721.html
18. Zentralinstitut für die kassenärztliche Versorgung. Patienten-Navi. Accessed May 15, 2020. https://www.116117.de/de/patienten-navi.php
19. Sana. Interaktiver Symptom-Checker, Accessed November 24, 2023. https://www.sana.de/unternehmen/digital/digitale-loesungen-im-einsatz/symptom-checker
20. AOK. AOK NAVIDA - Ihre persönliche Gesundheitsassisstentin. AOK. Accessed November 24, 2023. https://www.aok.de/pk/plus/app-navida-gesundheitsassistentin/
21. AXA. Symptom-Check Krankheitssymptome schnell einschätzen. Accessed November 24, 2023. https://www.axa.de/pk/gesundheit/s/symptom-check
22. Winn AN, Somai M, Fergestrom N, Crotty BH. Association of Use of Online Symptom Checkers With Patients' Plans for Seeking Care. *JAMA Netw Open*. 2019;2(12):e1918561. doi:10.1001/jamanetworkopen.2019.18561
23. Knitza J, Muehlensiepen F, Ignatyev Y, et al. Patient's Perception of Digital Symptom Assessment Technologies in Rheumatology: Results From a Multicentre Study. *Front Public Health*. 2022;10:844669. doi:10.3389/fpubh.2022.844669
24. Arellano Carmona K, Chittamuru D, Kravitz RL, Ramondt S, Ramirez AS. Health Information Seeking From an Intelligent Web-Based Symptom Checker: Cross-sectional Questionnaire Study. *J Med Internet Res*. 2022;24(8):e36322. doi:10.2196/36322
25. Verzantvoort NCM, Teunis T, Verheij TJM, van der Velden AW. Self-triage for acute primary care via a smartphone application: Practical, safe and efficient? *PLoS One*. 2018;13(6):e0199284. doi:10.1371/journal.pone.0199284
26. Pairon A, Philips H, Verhoeven V. A scoping review on the use and usefulness of online symptom checkers and triage systems: How to

proceed? *Front Med*.2023; 9:1040926. doi:10.3389/fmed.2022.104092
27. Miller S, Gilbert S, Virani V, Wicks P. Patients' Utilization and Perception of an Artificial Intelligence-Based Symptom Assessment and Advice Technology in a British Primary Care Waiting Room: Exploratory Pilot Study. *JMIR Hum Factors*. 2020;7(3):e19713. doi:10.2196/19713
28. Meyer AND, Giardina TD, Spitzmueller C, Shahid U, Scott TMT, Singh H. Patient Perspectives on the Usefulness of an Artificial Intelligence-Assisted Symptom Checker: Cross-Sectional Survey Study. *J Med Internet Res*. 2020;22(1):e14679. doi:10.2196/14679
29. Knitza J, Simon D, Lambrecht A, et al. Mobile Health Usage, Preferences, Barriers, and eHealth Literacy in Rheumatology: Patient Survey Study. *JMIR Mhealth Uhealth*. 2020;8(8):e19661. doi:10.2196/19661
30. Schaefer KE, Chen JY, Szalma JL, Hancock PA. A meta-analysis of factors influencing the development of trust in automation: Implications for understanding autonomy in future systems. *Hum factors*. 2016;58(3):377-400. doi: 10.1177/0018720816634228.
31. Aboueid S, Meyer S, Wallace JR, Mahajan S, Chaurasia A. Young Adults' Perspectives on the Use of Symptom Checkers for Self-Triage and Self-Diagnosis: Qualitative Study. *JMIR Public Health and Surveill*. 2021;7(1):e22637. doi:10.2196/22637
32. Woodcock C, Mittelstadt B, Busbridge D, Blank G. The Impact of Explanations on Layperson Trust in Artificial Intelligence-Driven Symptom Checker Apps: Experimental Study. *J Med Internet Res*. 2021;23(11):e29386. doi:10.2196/29386
33. Kopka M, Schmieding ML, Rieger T, Roesler E, Balzer F, Feufel MA. Determinants of Laypersons' Trust in Medical Decision Aids: Randomized Controlled Trial. *JMIR Hum Factors*. 2022;9(2):e35219. doi: 10.2196/35219
34. Nijland N, Cranen K, Boer H, van Gemert-Pijnen JE, Seydel ER. Patient use and compliance with medical advice delivered by a web-based triage system in primary care. *J Telemed Telecare*. 2010;16(1):8-11. doi:10.1258/jtt.2009.001004
35. Wetzel AJ, Koch R, Koch N, et al. 'Better see a doctor?' Status quo of symptom checker apps in Germany: A cross-sectional survey with a mixed-methods design (CHECK.APP). *Digit Health*. 2024;10:20552076241231555. doi:10.1177/20552076241231555
36. Wetzel AJ, Klemmt M, Müller R, Rieger MA, Joos S, Koch R. Only the anxious ones? Identifying characteristics of symptom checker app users: a cross-sectional survey. *BMC Med Inform Dec Mak*. 2024;24(1):21. doi: 10.1186/s12911-024-02430-5
37. Starcevic V, Berle D, Arnáez S. Recent insights into Cyberchondria. *Curr Psychiatry Rep*. 2020;22(11). doi:10.1007/s11920-020-01179-8
38. Arsenakis S, Chatton A, Penzenstadler L, et al. Unveiling the relationships between cyberchondria and psychopathological

symptoms*. J Psychiatr Res.* 2021;143:254-261. doi:10.1016/j.jpsychires.2021.11.020
39. Wetzel AJ, Müller R, Joos S, Koch R, Henking T, Haumann H. “Do I really not need to go to the doctor?” Unveiling usage patterns and predicting usage of Symptom Checker Apps: an explorative longitudinal study. *JMIR Preprints.* 2024. doi:10.2196/preprints.29202
40. Laux G, Kühlein T, Gutscher A, Szecsenyi J. *Versorgungsforschung in der Hausarztpraxis*. Urban und Vogel; 2010.
41. Abramowitz JS, Braddock A. *Psychological treatment of health anxiety and hypochondriasis: A biopsychosocial approach*. Hogrefe; 2008.
42. Bailer J, Kerstner T, Witthöft M, Diener C, Mier D, Rist F. Health anxiety and hypochondriasis in the light of DSM-5. *Anxiety Stress Coping.* 2016;29(2):219-239. doi:10.1080/10615806.2015.1053940
43. Ferguson E. A taxometric analysis of health anxiety. *Psychol Med.* 2009;39(2):277-285. doi:10.1017/S003329170800387X
44. First MB, France A, Pincus HA. *DSM-IV-TR guidebook*. American Psychiatric Publishing Inc.; 2004.
45. Fink P, Ornbol E, Christensen KS. The outcome of health anxiety in primary care. A two-year follow-up study on health care costs and self-rated health. *PLoS One.* 2010;5(3):e9873. doi:10.1371/journal.pone.0009873
46. Hannah K, Marie K, Olaf H, et al. The global economic burden of health anxiety/hypochondriasis- a systematic review. *BMC Public Health*. 2023;23(1):2237. doi:10.1186/s12889-023-17159-5
47. Dziegielewski S. DSM-IV-TR in action. 2nd ed. John Wiley & Sons; 2010.
48. American Psychiatric Association. *Diagnostic and Statistical Manual of Mental Disorders (DSM-5)*. 5th ed. American Psychiatric Publishing; 2013.
49. Sauer KS, Witthöft M. Krankheitsängste und Hypochondrische Störung. *PSYCH up2date*. 2020;14(2):151-166. doi:10.1055/a-0976-8998
50. Weck F, Richtberg S, MB Neng J. Epidemiology of hypochondriasis and health anxiety: comparison of different diagnostic criteria. *Curr Psychiatry Rev*. 2014;10(1):14-23. doi:10.2174/1573400510666140201000518
51. Bleichhardt G, Hiller W. Hypochondriasis and health anxiety in the German population. *Br J Health Psychol*. 2007;12(4):511-523. doi:10.1348/135910706X169359
52. Carleton RN. Into the unknown: A review and synthesis of contemporary models involving uncertainty. *J Anxiety Disord*. 2016;39:30-43. doi:10.1016/j.janxdis.2016.02.007.
53. McEvoy PM, Hyett MP, Shihata S, Price JE, Strachan L. The impact of methodological and measurement factors on transdiagnostic associations with intolerance of uncertainty: A meta-analysis. *Clin Psychol Rev*. 2019;73:101778. doi:10.1016/j.cpr.2019.101778

54. Elhamiasl M, Dehghani M, Heidari M, Vancleef LM, Khatibi A. Negative interpretation of ambiguous bodily symptoms among illness-anxious individuals: Exploring the role of developmental and maintenance constructs. *Front Psychiatry*. 2023;13:2812. doi:10.3389/fpsyt.2022.852881
55. Starcevic V. Cyberchondria: challenges of problematic online searches for health-related information*. Psychother Psychosom*. 2017;86(3):129-133. doi:10.1159/000453775
56. Mathes BM, Norr AM, Allan NP, Albanese BJ, Schmidt NB. Cyberchondria: Overlap with health anxiety and unique relations with impairment, quality of life, and service utilization. *Psychiatry Res*. 2018;261:204-211. doi:10.1016/j.psychres.2017.12.064
57. Fergus TA, Russell LH. Does cyberchondria overlap with health anxiety and obsessive–compulsive symptoms? An examination of latent structure and scale interrelations. *J Anxiety Disord.* 2016;38:88-94. doi:10.1016/j.janxdis.2016.01.007
58. Starcevic V, Baggio S, Berle D, Khazaal Y, Viswasam K. Cyberchondria and its relationships with related constructs: A network analysis. *Psychiatr Q*. 2019;90:491-505. doi:10.1007/s11126-019-09659-x
59. McElroy E, Shevlin M. The development and initial validation of the cyberchondria severity scale (CSS). *J Anxiety Disord*. 2014;28(2):259-265. doi:10.1016/j.janxdis.2013.12.007
60. Keller GL, Padala PR, Petty F. Clinical pearls to manage cyberchondriacs. *Prim Care Companion CNS Disord*. 2008;10(1):24295. doi:10.4088/PCC.v10n0107
61. Starcevic V, Berle D. Cyberchondria: towards a better understanding of excessive health-related Internet use. *Expert Rev Neurother*. 2013;13(2):205-213. doi:10.1586/ern.12.183
62. Fergus TA. Cyberchondria and intolerance of uncertainty: examining when individuals experience health anxiety in response to Internet searches for medical information. *Cyberpsychol Behav Soc Netw*. 2013;16(10):735-739. doi:10.1089/cyber.2012.0523
63. Hedman E, Lekander M, Ljótsson B, et al. Optimal Cut-Off Points on the Health Anxiety Inventory, Illness Attitude Scales and Whiteley Index to Identify Severe Health Anxiety. *PLoS One*. 2015;10(4):e0123412. doi:10.1371/journal.pone.0123412
64. Müller R, Klemmt M, Koch R, et al. “That’s just Future Medicine”-a qualitative study on users’ experiences of symptom checker apps. *BMC Med Ethics*. 2024;25(1):1-19. doi:10.1186/s12910-023-00783-1
65. Sankar R, Robinson L, Honey E, Freeston M. ‘We know intolerance of uncertainty is a transdiagnostic factor but we don’t know what it looks like in everyday life’: A systematic review of intolerance of uncertainty behaviours. *Clin Psychol Forum.* 2017;296:10-15. https://eprints.ncl.ac.uk/241951
66. Ceney A, Tolond S, Glowinski A, Marks B, Swift S, Palser T. Accuracy of online symptom checkers and the potential impact on

service utilisation. *PLoS One*. 2021;16(7):e0254088. doi:10.1371/journal.pone.0254088

67. Fraser HSF, Cohan G, Koehler C, et al. Evaluation of Diagnostic and Triage Accuracy and Usability of a Symptom Checker in an Emergency Department: Observational Study*. JMIR Mhealth Uhealth*. 2022;10(9):e38364. doi:10.2196/38364
68. Norbye AD, Abelsen B, Førde OH, Ringberg U. Health anxiety is an important driver of healthcare use. *BMC Health Serv Res.* 2022;22(1):138. doi:10.1186/s12913-021-07498-w

7. Declaration of Own Contributions

The dissertation was conducted at the Institute of General Practice and Interprofessional Care under the supervision of Prof. Dr. Stefanie Joos.

I declare that I have written the manuscripts by myself and that I have not used any additional references than those cited by me. Proof reading of selected paragraphs and sentences for language errors in spelling, grammar and word choice was performed by the Large Language Model ChatGPT by openAI and Marius Wetzel.

Tübingen, 28th May 2024

_____

Anna-Jasmin Wetzel

Declaration of contributions for the research project CHECK.APP

Project team:
Anna-Jasmin Wetzel (AJW), Roland Koch (RK), Christine Preiser (CP), Regina Müller (RM), Malte Klemmt (MK), Robert Ranisch (RR), Hans-Jörg Ehni (HJE), Urban Wiesing (UW), Monika Rieger (MR), Tanja Henking (TH), Stefanie Joos (SJ)

The overall conception of the CHECK.APP research project, on which the doctoral thesis was based, was carried out by SJ, RK, CP, RR, HJE, UW, MR and TH. AJW joined the team starting with the project began. AJW drafted and wrote the application for the ethics approval and got feedback from the whole project team.

Declaration of contributions for the publication Ethical, Legal, and Social Implications of Symptom Checker Apps in Primary Health Care (CHECK.APP): Protocol for an Interdisciplinary Mixed Methods Study

Authors:

Anna-Jasmin Wetzel (AJW), Roland Koch (RK), Christine Preiser (CP), Regina Müller (RM), Malte Klemmt (MK), Robert Ranisch (RR), Hans-Jörg Ehni (HJE), Urban Wiesing (UW), Monika Rieger (MR), Tanja Henking (TH), Stefanie Joos (SJ)

Journal:

JMIR Research Protocols

Contributions:

AJW and RK made equal contributions to drafting and writing the manuscript. AJW conducted the literature search for the theoretical background. Together, AJW and RK developed the theoretical framework and the discussion section. AJW authored the initial draft of the method section, while the selection of methods was a collaborative effort among AJW, SJ, and RK. Following feedback from the co-authors, AJW revised the manuscript and prepared it for publication. All authors participated in editing the draft and approved the final version of the manuscript.

Declaration of contributions for the publication "Better see a doctor?" Status Quo of Symptom Checker Apps in Germany: A Cross-Sectional Survey with Mixed-Method Design (CHECK.APP)

Authors:

Anna-Jasmin Wetzel (AJW), Roland Koch (RK), Nadine Koch (NK), Malte Klemmt (MK), Regina Müller (RM), Christine Preiser (CP), Monika Rieger (MR), Inka Rösel (IR), Robert Ranisch (RR), Hans-Jörg Ehni (HJE), Stefanie Joos (SJ)

Journal:

Sage Digital Health

Contributions:

AJW conducted literature searches to inform the development of the applied survey, research concept, methods, and theoretical background of the manuscript. The survey was developed by AJW and discussed with RK, and SJ. AJW was responsible for patient recruitment and data acquisition. The manuscript was conceptualized by AJW and SJ. AJW proposed the data selection, which was then discussed with SJ and RK. AJW conducted the data preparation, processing and analysis with support from RK. The initial draft of the manuscript was written by AJW. Results were interpreted by AJW and reviewed in interdisciplinary meetings with SJ, RK, MR, CP, RR, MK, and RM. Following co-author feedback, AJW revised and prepared the manuscript for publication. All authors participated in editing the draft and approved the final manuscript.

Declaration of contributions for the publication "Only the Anxious Ones?" Identifying Characteristics of Symptom Checker App Users: A Cross-sectional Survey.

Authors:

Anna-Jasmin Wetzel (AJW), Malte Klemmt (MK), Regina Müller (RM), Monika Rieger (MR), Stefanie Joos (SJ), Roland Koch (RK)

Journal:

BMC Medical Informatics and Decision Making

Contributions:

AJW conducted literature searches to support the development of the applied survey, research concept, methods, and theoretical background of the manuscript. AJW developed the survey with RK and SJ. AJW was also responsible for recruiting patients and acquiring data. The manuscript was initially conceptualized by AJW and RK, and further refined in collaboration with SJ. AJW proposed the data selection, which was then discussed and agreed upon with SJ and RK. AJW conducted the data preparation, processing and analysis. AJW authored the

first draft of the manuscript. Results were interpreted by AJW and discussed with SJ, RK, and MR during interdisciplinary meetings. After receiving feedback from the co-authors, AJW revised and prepared the manuscript for publication. All authors participated in editing the draft and approved the final manuscript.

Declaration of contributions for the publications "Do I really not need to go to the doctor?" Unveiling usage patterns and predicting usage of Symptom Checker Apps: an explorative longitudinal study.

Authors:
Anna-Jasmin Wetzel (AJW), Christine Preiser (CP), Regina Müller (RM), Stefanie Joos (SJ), Roland Koch (RK), Tanja Henking (TH), Hannah Haumann (HH)
Journal:
Journal of Medical Internet Research (JMIR, under Review)

Contributions:
AJW conducted literature reviews for developing the user diaries, as well as the research concept, methods, and theoretical framework of the manuscript. AJW drafted and piloted the user diaries, incorporating feedback from RK, RM, MK, and CP during interdisciplinary meetings. AJW was also responsible for patient recruitment and data acquisition. Together with HH, AJW conceptualized the manuscript. Data selection was initiated by AJW and further discussed with HH and CP. AJW conducted the data preparation, processing and analysis. The initial draft of the manuscript was written by AJW. Results were interpreted by AJW and reviewed with HH, RK, and CP. After incorporating feedback from co-authors, AJW revised and prepared the manuscript for publication. All authors participated in editing the draft and approved the final version.

8. Publications in Connection with the Thesis

Wetzel AJ, Koch R, Preiser C, Müller R, Klemmt M, Ranisch R, Joos S. Ethical, legal, and social implications of symptom checker apps in primary health care (CHECK. APP): protocol for an interdisciplinary mixed methods study. JMIR *Res Protoc*. 2022;11(5). doi:10.2196/34026

Radionova N, Ög E, **Wetzel** AJ, Rieger MA, Preiser C. Impacts of symptom checkers for laypersons' self-diagnosis on physicians in primary care: scoping review. *J Med Internet Res*. 2023;25. doi:10.2196/39219

Wetzel AJ, Koch R, Koch N, et al. 'Better see a doctor?' Status quo of symptom checker apps in Germany: A cross-sectional survey with a mixed-methods design (CHECK.APP). *Digit Health*. 2024;1 0:20552076241231555. doi:10.1177/20552076241231555

Wetzel A-, Klemmt M, Müller R, Rieger MA, Joos S, Koch R. Only the anxious ones? Identifying characteristics of symptom checker app users: a cross-sectional survey. *BMC Med Inform Dec Mak*. 2024;24(1):21. doi: 10.1186/s12911-024-02430-5

Wetzel AJ, Müller R, Joos S, Koch R, Henking T, Haumann H. "Do I really not need to go to the doctor?" Unveiling usage patterns and predicting usage of Symptom Checker Apps: an explorative longitudinal study. *JMIR Preprints.* 2024. doi:10.2196/preprints.29202

Koch R, Steffen MT, **Wetzel** AJ, Preiser C, Klemmt M, Ehni H-J, Mueller R, Joos S. An interface between ehealth, health literacy and health-related behavior: Qualitative interview study on laypersons experiences with a mobile symptom checker app. *JMIR Preprints*. doi:10.2196/preprints.60647.

9. Acknowledgments

I would like to express my gratitude to Prof. Dr. Joos for her trust, freedom, and guidance over the past few years. My dissertation would not have been possible without these factors and our respectful collaboration. I am thankful to Dr. Christine Preiser, Dr. Hannah Haumann, and Dr. David Häske for their sound scientific advice, open ears, and especially for their encouragement. Young scientists need inspiring and modern role models like all of you. Additionally, I wish to thank my family. My husband, Marius, who has always supported me wholeheartedly, shared my excitement at all times, brainstormed ideas, helped with language revision, and stood by me unconditionally. My father, Helmut, who enabled my academic career and was always ready to assist wherever possible. To my brothers, who encouraged me to learn new things and believe in myself. I also want to thank my young daughter, Jara, who has taught me the importance of a fresh perspective and how effective time management can be.