<u>Original Paper</u>

Impact of a Symptom Checker App on Patient-Physician Interaction Among Self-Referred Walk-In Patients in the Emergency Department: Multicenter, Parallel-Group, Randomized, Controlled Trial

Malte L Schmieding[1], MBI, Dr med; Marvin Kopka[1,2], MSc, Dr rer med; Myrto Bolanaki[3], Dr med; Hendrik Napierala[4], MSc, Dr med; Maria B Altendorf[3], PhD; Doreen Kuschick[4], MSc, Dr rer med; Sophie K Piper[1,5], Dr rer nat; Lennart Scatturin[1], MSc; Konrad Schmidt[4], Prof Dr Med; Claudia Schorr[4], MSc; Alica Thissen[1], Dr phil; Cornelia Wäscher[4], BSc, MPH; Christoph Heintze[4], MA, MPH, Prof Dr Med; Martin Möckel[3], Prof Dr Med; Felix Balzer[1], Dr rer nat, Prof Dr Med; Anna Slagman[3], MSc, Prof Dr

[1]Institute of Medical Informatics, Charité – Universitätsmedizin Berlin, corporate member of Freie Universität Berlin and Humboldt-Universität zu Berlin, Berlin, Germany

[2]Division of Ergonomics, Department of Psychology and Ergonomics, Technische Universität Berlin, Berlin, Germany

[3]Emergency and Acute Medicine, Campus Virchow-Klinikum and Campus Charité Mitte, Charité – Universitätsmedizin Berlin, corporate member of Freie Universität Berlin and Humboldt-Universität zu Berlin, Berlin, Germany

[4]Institute of General Practice and Family Medicine, Charité – Universitätsmedizin Berlin, corporate member of Freie Universität Berlin and Humboldt-Universität zu Berlin, Berlin, Germany

[5]Institute of Biometry and Clinical Epidemiology, Charité – Universitätsmedizin Berlin, corporate member of Freie Universität Berlin and Humboldt-Universität, Berlin, Germany

Corresponding Author:
Malte L Schmieding, MBI, Dr med
Institute of Medical Informatics
Charité – Universitätsmedizin Berlin, corporate member of Freie Universität Berlin and Humboldt-Universität zu Berlin
Charitéplatz 1
Berlin, 10117
Germany
Phone: 49 450 570425
Email: malte.schmieding@charite.de

Abstract

Background: Symptom checker apps (SCAs) are layperson-facing tools that advise on whether and where to seek care, or possible diagnoses. Previous research has primarily focused on evaluating the accuracy, safety, and usability of their recommendations. However, studies examining SCAs' impact on clinical care, including the patient-physician interaction and satisfaction with care, remain scarce.

Objective: This study aims to evaluate the effects of an SCA on satisfaction with the patient-physician interaction in acute care settings. Additionally, we examined its influence on patients' anxiety and trust in the treating physician.

Methods: This parallel-group, randomized controlled trial was conducted at 2 emergency departments of an academic medical center and an emergency practice in Berlin, Germany. Low-acuity patients seeking care at these sites were randomly assigned to either self-assess their health complaints using a widely available commercial SCA (Ada Health) before their first encounter with the treating physician or receive usual care. The primary endpoint was patients' satisfaction with the patient-physician interaction, measured by the Patient Satisfaction Questionnaire (PSQ). The secondary outcomes were patients' satisfaction with care, their anxiety levels, and physicians' satisfaction with the patient-physician interaction. We used linear mixed models to assess the statistical significance of primary and secondary outcomes. Exploratory descriptive analyses examined patients' and physicians' perceptions of the SCA's utility and the frequency of patients questioning their physician's authority.

Results: Between April 11, 2022, and January 25, 2023, we approached 665 patients. A total of 363 patients were included in the intention-to-treat analysis of the primary outcome (intervention: n=173, control: n=190). PSQ scores in the intervention group

were similar to those in the control group (mean 78.5, SD 20.0 vs mean 80.8, SD 19.6; estimated difference –2.4, 95% CI –6.3 to 1.1, P=.24). Secondary outcomes, including patients' and physicians' satisfaction with care and patient anxiety, showed no significant group differences (all P>.05). Patients in the intervention group were more likely to report that the SCA had a beneficial (66/164, 40.2%) rather than a detrimental (3/164, 1.8%) impact on the patient-physician interaction, with most reporting no effect (95/164, 57.9%). Similar patterns were observed regarding the SCA's perceived effect on care. In both groups, physicians rarely reported that their authority had been questioned by a patient (intervention: 2/188, 1.1%; control: 4/184, 2.2%). While physicians more often found the SCA helpful rather than unhelpful, the majority indicated it was neither helpful nor unhelpful for the encounter.

Conclusions: We found no evidence that the SCA improved satisfaction with the patient-physician interaction or care in an acute care setting. By contrast, both patients and their treating physicians predominantly described the SCA's impact as beneficial. Our study did not identify negative effects of SCA use commonly reported in the literature, such as increased anxiety or diminished trust in health care professionals.

Trial Registration: German Clinical Trial Register DRKS00028598; https://drks.de/search/en/trial/DRKS00028598/entails

International Registered Report Identifier (IRRID): RR2-10.1186/s13063-022-06688-w

KEYWORDS

Introduction

It has become common for the general population to seek health-related information online. In the European Union, more than 1 in 2 citizens searched for health-related information online in the 3 months preceding the survey [1]. Similarly, 74.4% of US adults consulted the internet first when seeking health information during their most recent inquiry [2]. In a German panel study, a fifth of respondents identified the internet as their primary source of health information [3]. In particular, seeking health-related information online is common before consulting medical services among low-urgency acute care patients [4,5]. While qualitative studies highlight concerns from both patients [6-8] and health care professionals [9-12] regarding online information exacerbating patient anxiety [13] and undermining the patient-physician relationship, 2 quantitative observational studies investigating the effects of online information seeking showed positive effects on the perceived quality of care received and the patient-physician interaction [5,14]. However, results from an interventional study measuring these metrics as secondary outcomes did not provide evidence for such a positive effect [4].

One particular source of online health information is a symptom checker app (SCA). These consumer apps offer suggestions on potential diagnoses or an urgency assessment based on the self-reported signs and symptoms entered by users. SCAs face similar concerns to web-based health information–seeking behavior in general, as described above, particularly regarding the induction of anxiety [15-17]. Studies estimate that the proportion of SCA users in the German population ranges between 6% and 13% [16,18,19]. Some national health care services [20,21], health care systems [22,23], hospital networks [24], and insurance companies [25] have already integrated SCAs into their service pathways.

So far, research studies have primarily focused on assessing the accuracy and safety of SCA advice [26-38], their (potential) impact on patient journeys and resource allocation [28,39-43], and users' experiences and expectations [6,7,17,30,38,41,44], but not on SCAs' actual impact on patient-centered, clinically relevant outcomes. Four recent reviews concluded that more research on SCAs' impacts in real-life settings is needed to assess their utility [45-48]. To address this research gap, we conducted a multicenter randomized controlled trial evaluating the effect of SCA usage in an acute care setting. We focused on patient-physician interaction, satisfaction with care, and users' anxiety, as these areas frequently feature in discussions about the impact of SCAs and online health information on care [4,5,9,10,17,49]. We excluded questions regarding the SCA's utility for improving patient allocation. The primary hypothesis was that the intervention would enhance patients' satisfaction with their interaction with the treating physician. Secondary hypotheses included improvements in patients' satisfaction with the health care received and their anxiety levels.

Methods

Study Design, Setting, and Participants

We conducted a multicenter, controlled parallel-group trial with balanced randomization at 3 study sites in Berlin, Germany. Two study sites were the emergency departments (EDs) of a large tertiary care university hospital: CCM (Charité – Universitätsmedizin Berlin, Campus Mitte in Berlin-Mitte) and CVK (Campus Virchow-Klinikum in Berlin-Wedding). Both EDs provide a wide spectrum of nonpediatric care, handling approximately 50,000 patient encounters annually, and operate an adjunct acute medical admissions ward. The third study site was an emergency practice operated by Berlin's Association of Statutory Health Insurance Physicians (Kassenärztliche Vereinigung Berlin), located adjacent to the ED of a local hospital (Jüdisches Krankenhaus Berlin in

Berlin-Gesundbrunnen [JKB]). This outpatient clinic provides urgent care outside regular office hours for approximately 4000 walk-in patients per year. It is typically staffed by 1 specialist physician and 1 medical assistant. The emergency practice serves self-referred patients who do not require treatment in the hospital-run ED located in the same building. This stratification of emergency care is designed to ensure that inpatient resources are reserved for urgently triaged patients. Key inclusion criteria were self-referred walk-in patients aged 18 years or older with sufficient German or English language proficiency, the ability to provide informed consent, and a treatment urgency rating of yellow, green, or blue according to the Manchester Triage System (ie, MTS 3-5, respectively) as assigned by the triage nurse. Exclusion criteria were patients treated without waiting time; patients whose chief complaint was already known to them; patients requiring isolation; patients unable to handle a tablet computer, as determined by either their own assessment or that of the study personnel; and patients who had already consulted an SCA for their current complaints before seeking care.

Randomization and Masking

Participants were randomly assigned to either use the SCA before their first encounter with a treating physician or receive care as usual (1:1 ratio). Balanced block randomization with variable block lengths of 8, 10, and 12 was used. The recruiting study personnel were blinded to block size. Each trial site received its own allocation sequence, which was stored in sequentially numbered, opaque, and sealed envelopes. The allocation sequence was generated by the Institute of Medical Informatics using the R package blockrand (R Foundation) [50] by a researcher (MLS) who was not actively involved in recruitment. Allocation was concealed until the point of randomization, which occurred immediately after the patient consented to participate in the trial. Because of the nature of the intervention, participants, study personnel, and health care providers were not masked to group assignment after randomization.

Procedures

Following initial triage, patients underwent eligibility screening conducted by study personnel (study nurse, student research assistant, or study physician). Upon providing informed consent, they were randomly assigned to either the control or intervention group (see details on randomization below). Before their first encounter with the treating physician, all participants completed a baseline survey assessing baseline anxiety, prior SCA use, and affinity for technology interaction [51]. A survey on sociodemographic and other variables that do not change during a patient's stay—such as age, sex, native language, country of residence, level of education, net household income, frequency of internet, tablet, and smartphone use, self-perceived health and chronic morbidity [52], self-efficacy [53], and eHealth literacy [54]—was administered at a suitable time during their visit, either before or after their encounter with the treating physician (sociodemographic survey). Additionally, after seeing the treating physician—or in exceptional cases, within 72 hours thereafter—all participants were asked to complete a postencounter survey. Participants in the intervention group completed the self-assessment of their symptoms using the SCA after taking the baseline survey but before their first encounter with the treating physician. Following SCA use and before their consultation with a physician, they completed a post-SCA survey assessing their experience with the SCA and their level of anxiety. All surveys were administered via a tablet computer, with study personnel providing instructions on its use.

Study personnel ensured that a printout of the SCA summary report was available to treating physicians at the time of their first encounter with patients in the intervention group. However, to avoid interfering with the patient-physician relationship and care provision, study personnel neither encouraged nor discouraged physicians from engaging with the summary report. All participating physicians were aware that such reports would be available for patients in the intervention group. After providing care, the participating physicians completed a paper-based survey (physician-sided Patient Satisfaction Questionnaire [PSQ]), which assessed their satisfaction with the care provided, whether the patient questioned their authority, and their appraisal of the SCA's helpfulness and impact on the patient encounter.

At the JKB trial site, all patients presenting on recruiting days were screened for eligibility. However, due to the higher patient volume at CCM and CVK, it was not feasible to screen all patients for eligibility at these sites.

The SCA used in the intervention (2024 Ada Health GmbH [55]) was developed independently of the researchers involved in this trial. We selected this particular commercial SCA from among the many available based on existing literature regarding its reported diagnostic and triage accuracy, the safety of its advice, its usability, the breadth of conditions and chief complaints covered [28,33,44,56], and its availability in both German and English. The SCA requires users to provide basic demographic information (such as age and sex), past medical history (including smoking behavior and prior diagnoses), and details about their current medical complaints. In this initial step, users can select an unlimited number of symptoms [57]. They then provide additional information by answering a series of closed questions with binary or multiple-choice answer options, presented by the SCA in a conversational format. These "conversations" typically last about 6-8 minutes [29,56,58]. The SCA then provides an assessment of the urgency of the complaints and suggests 1-5 probable causes (diagnostic suggestions), illustrated using a Sankey diagram. During the trial period, the SCA did not inquire about users' intent to seek care, their own urgency assessment of their complaints, or the diagnoses they suspected. The SCA's report also summarizes the findings that the user affirmed or denied. The SCA generates its suggestions using a medical knowledge base that hard-codes libraries of signs, symptoms, and diagnoses, along with their relationships, and applies a Bayesian network algorithm [59]. The company behind Ada Health describes this algorithm as "artificial intelligence" [60]. Further details of the SCA's underlying algorithm are not publicly available. According to a 2024 study [16], Ada Health is the second most frequently used SCA in Germany. The developers also report that it is widely used in other countries, including Australia [61].

In the care-as-usual group, patients used the tablet only to complete the questionnaires (baseline, sociodemographic, and postencounter surveys) and did not have access to the SCA for self-assessing their symptoms. Patients in both groups were not prohibited from searching for online health information or using an SCA on their own devices during the study.

Study personnel closely observed patients for any signs of discomfort during the study procedures. In both groups, clinical routine care was conducted as usual. Study procedures were interrupted as needed for clinical interventions or other necessary reasons.

Outcomes

Primary Outcome

The primary outcome was participants' satisfaction with their interaction with the treating physician, assessed using the PSQ [62,63]. This instrument consists of visual analog scales ranging from 0 to 100 for each of the 5 items. A participant's overall satisfaction is defined as the average score across all 5 items, with higher values indicating greater satisfaction. Participants' responses on the primary outcome were collected in the postencounter survey before discharge or, in exceptional cases, within 72 hours of discharge if they left the trial site without completing the questionnaire. We considered a mean difference of 5 points (on a 0-100 scale) between treatment groups to be clinically relevant.

Secondary and Exploratory Outcomes

Participants' satisfaction with the care received was measured using the 8-item Fragebogen zur Patientenzufriedenheit (ZUF-8) [64], the German version of the 8-item Client Satisfaction Questionnaire (CSQ-8) [65]. The ZUF-8 scale ranges from 8 to 32 points, with higher values indicating greater satisfaction. As some participants did not respond to all items of the ZUF-8 (each rated on a scale from 1 to 4), we calculated the average value of the items they answered. Therefore, we report our results for the ZUF-8 on a scale from 1 to 4. This deviates from the protocol, which did not include modifications to account for missing values. We expected a 1-point difference in care satisfaction between the intervention and control groups on the original ZUF-8 scale, corresponding to 0.125 points on the 1-4 scale.

Participants' anxiety was measured up to 3 times during the trial: initially after recruitment (baseline), after using the SCA for participants in the intervention group, and finally after the patient-physician interaction. We used a visual analog scale ranging from 0 to 100, with higher values indicating greater anxiety [66]. Physicians' satisfaction with the patient-physician interaction was measured using the physician version of the PSQ [62,63], which rephrases the PSQ items from the physician's perspective.

As additional exploratory outcomes, we report participants' perceived effect of the SCA on the patient-physician interaction and patient care (intervention group only; two 5-point Likert scales), the physicians' assessment of the helpfulness of the SCA report (intervention group only; five 3-point Likert scales), and physicians' satisfaction with the care they delivered, including overall satisfaction, time to diagnosis, and patient length of stay (measured on a visual analog scale ranging from 0 to 100).

Sample Size Calculation

The trial group made a reasoned choice that a mean difference of at least 5 points in the PSQ score after the physician encounter between the intervention and control groups would be clinically relevant. In the literature, we found an SD of the patient-facing PSQ ranging from 14 to 17 points [63]. Assuming a standardized mean difference of 0.3 and equal variance, and considering a 2-sided α of .05, a power of 0.80, and a medium-sized dropout rate of 20%, a total of 440 patients were needed (ie, 220 in each trial arm). For the purpose of sample size calculation, we conservatively used a 2-sample (paired) t test for independent groups. However, our a priori planned analysis of the primary endpoint involves a mixed-effects model, which accounts for the clustered data structure (details below).

Statistical Analyses

Descriptive statistics are presented as mean with SD, median with IQR, or frequency and proportion, depending on the scale and distribution.

We conducted our primary analysis based on the modified intention-to-treat principle, including all randomly assigned patients who provided responses to at least one item related to the respective primary or secondary outcome measure in the postencounter survey. Subgroup analyses by study site were conducted as preplanned.

We analyzed the primary outcome—patients' satisfaction with the patient-physician interaction (measured by the PSQ)—using a linear mixed model with intervention as a fixed effect and study site as a random effect. Only participants who responded to at least one PSQ item were included. We report group means and SDs, the linear mixed model estimator, 95% CIs, and P values. Statistical analyses were conducted using R (version 4.4.0; R Foundation) [67], with data cleaning performed using tidyverse [68]. CIs were bootstrapped, and P values were calculated using the R packages lme4 [69] and parameters [70]. For secondary outcomes (patients' satisfaction with the care received, change in anxiety levels after SCA use, the proportion of participants more anxious after the physician encounter than at baseline, and physicians' satisfaction with patient-physician interaction), our analyses followed the same approach as for the primary outcome. As planned, we conducted no adjustments of P values due to multiplicity. To address missing data in primary and secondary outcomes, multiple imputation techniques were used as sensitivity analysis (see Table S3 in Multimedia Appendix 1). Regarding exploratory outcomes, we report descriptive statistics only.

The researchers statistically assessing the primary outcome were blinded to group assignment. Although not described in the previously published study protocol, we imputed missing data for primary and secondary endpoints as a sensitivity analysis using the R package mice [71].

Ethical Considerations

All participants provided written informed consent. We made consent forms and participant-facing information leaflets on the study available in English and German. The information leaflet included information on the study's intended purpose, design, procedure, the responsible persons, how and by whom their personal data were processed for the purposes of the study, and the participants' rights including rights stemming from the European Union's General Data Protection Regulation. The trial, the consent forms, and information leaflets were approved by the Institutional Review Board at Charité – Universitätsmedizin Berlin (reference number: EA2/284/21). We put in place technical and organizational measures to adhere to the European Union's General Data Protection Regulation with the support of the Charité – Universitätsmedizin Berlin's Clinical Trial Office. These measures included pseudonymizing data before conducting analyses, limiting access to all trial data, and storing and processing potentially reidentifiable trial data on servers owned and maintained by Charité – Universitätsmedizin Berlin. To render data reported in this paper anonymous, we only report aggregate statistics. The trial was prospectively registered in the German Clinical Trials Register (DRKS-ID: DRKS00028598). The protocol was previously published [72]. This manuscript follows the CONSORT (Consolidated Standards of Reporting Trials) statement [73]. Each participating patient received a €10 (US $10.8) gift voucher redeemable at over 300 brands and e-commerce platforms. Treating physicians received a €5 (US $5.42) gift voucher for each survey they completed on an encounter with a patient enrolled in the trial.

Role of the Funding Source

The study funder had no role in the study design, data collection, data analysis, data interpretation, report writing, or the decision to submit for publication. Similarly, the developer of the SCA, Ada Health, was not a study sponsor and had no involvement in any of these aspects.

Results

Recruitment and Participant Characteristics

Recruitment took place from April 11, 2022, to January 25, 2023. The recruitment phase was extended beyond the planned 6 months to accommodate staff shortages and COVID-19–related restrictions. Recruitment concluded upon reaching the predetermined sample size. A total of 442 participants were enrolled, with 220 randomly assigned to the intervention group and 222 to the control group. Two eligible participants in the control group dropped out immediately after randomization—due to organizational issues (n=1) and a medical condition (n=1). Three participants assigned to the intervention group withdrew their consent, and 5 additional participants were lost due to organizational issues. Of the 434 included participants, 188 (43.3%) identified as male, 220 (50.7%) as female, and 4 (0.9%) as diverse (with 22 participants not indicating their sex). The median age was 33 (IQR 26-45) years. Full details of group characteristics at baseline are provided in Table 1. The baseline survey was completed by 426 participants (206 in the intervention group and 220 in the control group). Of these, 363 provided sufficient data to assess the primary outcome (173 in the intervention group and 190 in the control group), with all but 2 (both in the intervention group) completing all surveys. Figure 1 presents the number of participants who completed the surveys over time.

Table S1 in Multimedia Appendix 1 compares the sex and age distributions of enrolled patients with those of all patients who presented at the trial sites during the study period. See Multimedia Appendix 2 for the CONSORT checklist.

Table 1. Baseline characteristics[a].

Characteristics	Control (n=222)	Intervention (n=212)
Age (years), median (IQR)	36 (28-47)	30 (25-41)
Sex, n (%)		
Male	96 (43.2)	92 (43.4)
Female	122 (55.0)	98 (46.2)
Other	2 (0.9)	2 (0.9)
NA[b]	2 (0.9)	20 (9.4)
Questionnaire completed in English, n (%)	49 (22.1)	54 (25.5)
Native language, n (%)		
German	120 (54.1)	103 (48.6)
English	20 (9.0)	20 (9.4)
Other	78 (35.1)	68 (32.1)
NA	4 (1.8)	21 (9.9)
Residence, n (%)		
Currently residing in Germany	209 (94.1)	175 (82.5)
Currently residing outside of Germany	10 (4.5)	16 (7.5)
NA	3 (1.4)	21 (9.9)
Education, n (%)		
Student	6 (2.7)	7 (3.3)
Basic education	1 (0.5)	0 (0)
Lower secondary education after year 9 or 10	13 (5.9)	12 (5.7)
Lower secondary education until year 10	38 (17.1)	25 (11.8)
High school degree for university entrance	135 (60.8)	124 (58.5)
Other school leaving certificate (eg, awarded abroad)	21 (9.5)	23 (10.8)
NA	8 (3.6)	21 (9.9)
Professional qualification, n (%)		
No qualification; still undergoing professional training (eg, student, trainee, in prevocational training, intern)	28 (12.6)	26 (12.3)
No professional qualification and not undergoing training	12 (5.4)	13 (6.1)
Apprenticeship (receiving vocational training from a company)	21 (9.5)	10 (4.7)
Training at a vocational school or a commercial school (combined vocational and academic education)	32 (14.4)	17 (8.0)
Technical college (eg, guild school, school for technicians, or vocational or professional academy)	14 (6.3)	11 (5.2)
University of applied sciences, engineering college, advanced technical college	18 (8.1)	16 (7.5)
University or college	82 (36.9)	87 (41.0)
Other formal qualification (eg, acquired abroad)	5 (2.3)	8 (3.8)
NA	10 (4.5)	24 (11.3)
Net household income[c], n (%)		
Less than €1800	53 (23.9)	61 (28.8)
€1801 to €2800	50 (22.5)	41 (19.3)
€2801 to €4000	44 (19.8)	25 (11.8)
More than €4000	39 (17.6)	43 (20.3)

Characteristics	Control (n=222)	Intervention (n=212)
NA	36 (16.2)	42 (19.8)
Anxiety (0-100), median (IQR)	73 (60-86)	72 (60-80)
Triage category, n (%)		
3	85 (38.3)	67 (31.6)
4	104 (46.8)	110 (51.9)
5	3 (1.4)	1 (0.5)
NA	30 (13.5)	34 (16.0)
Self-efficacy (ASKU[d]), median (IQR)	4.0 (3.3-4.3)	4.0 (3.3-4.3)
eHealth literacy (eHEALS[e]), median (IQR)	3.5 (3.0-3.9)	3.6 (3.1-4.1)
Self-perceived health, n (%)		
Very good	38 (17.1)	43 (20.3)
Good	110 (49.5)	83 (39.2)
Fair	54 (24.3)	47 (22.2)
Bad	14 (6.3)	18 (8.5)
Very bad	4 (1.8)	1 (0.5)
NA	2 (0.9)	20 (9.4)
Chronic morbidity, n (%)		
Yes	99 (44.6)	74 (34.9)
No	115 (51.8)	115 (54.2)
NA	8 (3.6)	23 (10.8)
Internet usage, n (%)		
Daily	201 (90.5)	178 (84.0)
Several times a week	15 (6.8)	9 (4.2)
Several times a month	4 (1.8)	2 (0.9)
Never	0 (0)	3 (1.4)
NA	2 (0.9)	20 (9.4)
Tablet usage, n (%)		
Daily	37 (16.7)	33 (15.6)
Several times a week	32 (14.4)	22 (10.4)
Several times a month	24 (10.8)	21 (9.9)
Several times a year	36 (16.2)	31 (14.6)
Never	86 (38.7)	83 (39.2)
NA	7 (3.2)	22 (10.4)
Previously used SCAs[f]	12 (5.4)	14 (6.6)

[a]We conducted a multicenter, controlled parallel-group trial at 3 study sites in Berlin, Germany. Between April 2022 and January 2023, we recruited self-referred adult walk-in patients presenting with an acute, undiagnosed chief complaint. The intervention group self-assessed their complaints using an SCA before their first encounter with a treating physician, while the control group received care as usual. After the patient-physician encounter, we surveyed both patients and physicians regarding the patient-physician interaction and their satisfaction with care.

[b]NA refers to both missing responses and respondents who indicate their preference not to provide a response. The higher dropout rate in the intervention group yields a greater proportion of NAs relative to the control group.

[c]€1=US $1.08.

[d]ASKU: Allgemeine Selbstwirksamkeit Kurzskala.

[e]eHEALS: eHealth Literacy Scale.

[f]SCA: symptom checker app.

Figure 1. Trial profile outlining the total number of patients screened, randomized, completing study surveys at multiple time points during their stay, and included in the primary end point analysis. We conducted a multicenter, controlled, parallel-group trial at 3 study sites in Berlin, Germany. Between April 2022 and January 2023, we recruited self-referred adult walk-in patients presenting with an acute, undiagnosed chief complaint. The intervention group self-assessed their complaints using an SCA before their first encounter with a treating physician, while the control group received usual care. After the patient-physician encounter, we surveyed both patients and physicians on the patient-physician interaction and their satisfaction with care. SCA: symptom checker app.
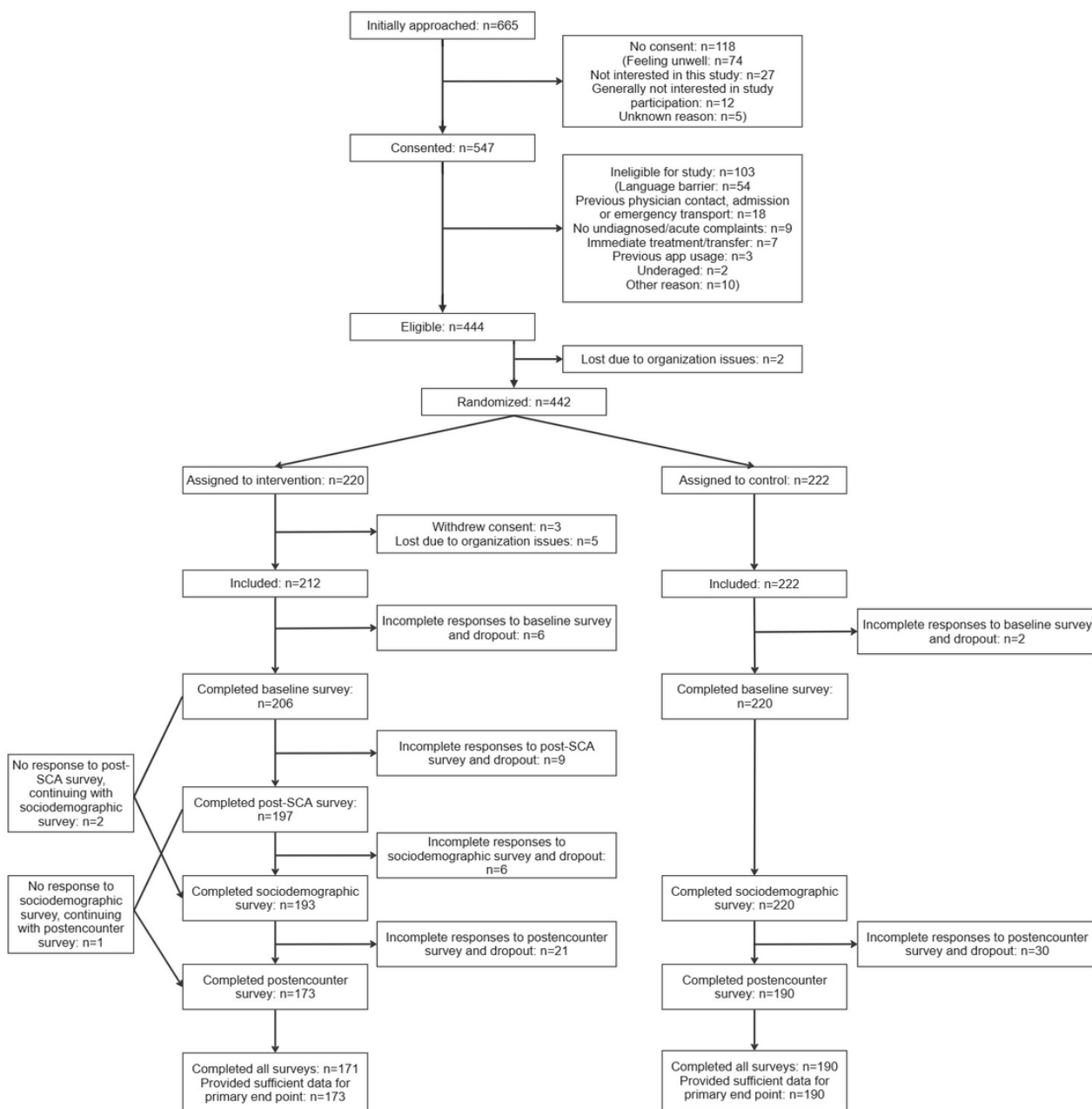
Primary Outcome

Across all 3 recruitment sites, participants' mean PSQ scores were close to 80 (on a scale from 0 to 100) in both groups (Table 2). The linear mixed-effects model showed no significant fixed effect for the group (estimate –2.4, 95% CI –6.3 to 1.1, reference: control group, *P*=.24). While patient-reported PSQ scores were similar between the control and intervention groups at 2 recruitment sites (CCM and JKB), participants from the CVK site in the intervention group reported, on average, 11 points lower satisfaction with the patient-physician interaction compared with their controls (see Table S2 in Multimedia Appendix 1). The use of different imputation methods showed no significant effects of missing data (see Table S3 in Multimedia Appendix 1).

Table 2. Primary and secondary endpoints according to each study group in the modified intention-to-treat population[a].

Endpoints	Control	Intervention	*P* value
Patient satisfaction with patient-physician interaction (patient-sided PSQ[b])			.24
Descriptive, mean (SD); n	80.8 (19.6); 190	78.5 (20); 173	
Estimate for the fixed effect of the study group in the linear mixed model (intervention to control group), 95% CI	N/A[c]	–2.4 (–6.3 to 1.1)	
Patient satisfaction with care (ZUF-8[d])			.27
Descriptive, mean (SD); n	2.6 (0.2); 190	2.6 (0.2); 173	
Estimate for the fixed effect of the study group in the linear mixed model (intervention to control group), 95% CI	N/A	0.02 (–0.02 to 0.06)	
Change in anxiety level, before SCA[e] use to after			.96
Descriptive, mean (SD); n	N/A	–1.7 (13.5); 199	
Estimate for the fixed effect of the study group in the linear mixed model, 95% CI	N/A	–0.1 (–5.0 to 4.5)	
Participants more anxious after the physician encounter than at baseline			.93
n/N (%)	39/191 (20.4)	36/173 (20.8)	
Estimate for the fixed effect of the study group in the generalized linear mixed model, 95% CI	N/A	0.0 (–0.5 to 0.6)	
Physician satisfaction with patient-physician interaction (physician-sided PSQ)			.08
Descriptive, mean (SD); n	76.3 (14.9); 203	73.7 (15.3); 191	
Estimate for the fixed effect of the study group in the linear mixed model (intervention to control group), 95% CI	N/A	–2.7 (–5.4 to 0.5)	

[a]Patient anxiety was assessed at baseline, after using the SCA (intervention group only), and after the physician encounter using a visual analog scale ranging from 0 to 100, with lower values indicating less anxiety. We conducted a multicenter, controlled parallel-group trial at 3 study sites in Berlin, Germany. Between April 2022 and January 2023, we recruited self-referred adult walk-in patients presenting with an acute, undiagnosed chief complaint. The intervention group self-assessed their complaints using an SCA before their first encounter with a treating physician, while the control group received care as usual. After the patient-physician encounter, we surveyed both patients and physicians regarding the patient-physician interaction and their satisfaction with care. Additionally, we assessed patients' anxiety about their symptoms at multiple time points during the trial—before and after the physician encounter, and, in the intervention group, also after using the SCA.

[b]PSQ: Patient Satisfaction Questionnaire.

[c]N/A: not applicable.

[d]ZUF-8: Fragebogen zur Patientenzufriedenheit.

[e]SCA: symptom checker app.

Secondary Outcomes

Satisfaction With Care (ZUF-8)

On average, participants in the intervention and control groups reported similar ZUF-8 average scores (control group: mean 2.6, SD 0.2, n=190; intervention group: mean 2.6, SD 0.2, n=173). The estimated group difference (0.02, 95% CI –0.02 to 0.06, *P*=.27) was lower than the hypothesized 0.125.

Anxiety Induced by SCA Usage

On average, patients reported slightly lower anxiety levels immediately after using the SCA (66.7 vs 64.5, ie, –1.7 points on a scale from 0 to 100, n=199), though this decrease was not statistically significant (estimate –0.1, 95% CI –5.0 to 4.5, *P*=.96). Approximately one-third of patients (70/199, 35.2%) in the intervention group reported increased anxiety after the SCA assessment, with about half of them (36/70) experiencing an increase of more than 5 points. Meanwhile, one-quarter (52/199, 26.1%) of patients in the intervention group reported a decrease in anxiety by more than 5 points. In both groups, one-fifth of participants reported higher anxiety levels after the physician encounter compared with their baseline level (see Table 2).

Treating Physicians' Satisfaction With Interaction (Physician-Sided PSQ)

Participating physicians completed the physician-facing PSQ in 40 of the 434 (9.2%) cases, with scores missing for 19 out of 222 patients in the control group and 21 out of 212 patients in the intervention group. On average, physicians reported lower mean PSQ scores for patients in the intervention group (73.7, SD 15.3) compared with those in the control group (76.3, SD 14.9). This group difference was less than the 5 points (on a scale from 0 to 100) deemed relevant in the study protocol and was not statistically significant (95% CI –5.4 to 0.5, *P*=.08). In 2 of the 4 imputation methods applied, mean differences reached significance (see Table S3 in Multimedia Appendix 1). These differences remained below the predefined 5-point group difference deemed clinically relevant. For all remaining secondary outcomes, none of the 4 imputation approaches

yielded statistically significant differences (see Table S3 in Multimedia Appendix 1).

Differences in primary and secondary outcome measures between the intervention and control groups did not reach statistical significance, even when analyzing only patients in the intervention group whose treating physician reviewed the SCA report (see Table S7 in Multimedia Appendix 1).

Further Exploratory Analyses

Patient-Reported Effects of the SCA

When asked about the perceived effect of the SCA on the patient-physician interaction, more than half of the participants in the intervention group reported no effect (95/164, 57.9%). More participants reported a (rather) positive influence (66/164, 40.2%) than a (rather) negative effect (3/164, 1.8%). Similar results were observed regarding the perceived effect on the care received (see Table S4 in Multimedia Appendix 1). These findings remained consistent even when considering only cases where the physician indicated having reviewed the SCA report (data not shown).

Effects of the SCA Based on Physician-Reported Outcomes

Physicians reported having taken notice of the SCA's summary report and recommendations for the majority of patients in the intervention group (112/187, 59.9%). Most physicians indicated that the SCA was neither helpful nor unhelpful for the 5 prespecified tasks (see Table S5 in Multimedia Appendix 1). However, for all 5 tasks, the SCA was rated as (rather) helpful more often than (rather) unhelpful. The SCA was considered most helpful for history taking, diagnosis, and conveying information to the patient. This finding remains unchanged when considering only cases in which the treating physician indicated having seen the SCA's report (data not shown).

Physicians rated their satisfaction with the patient care they provided, the adequacy of time to diagnosis, and the overall length of patient stay as similar across both trial groups (see Table S6 in Multimedia Appendix 1). In both groups, treating physicians reported only a few instances of patients questioning their authority (intervention group: 2/188, 1.1%; control group: 4/184, 2.2%).

Discussion

Principal Findings

The AkuSym study was the first RCT to examine the impact of using an SCA before contact with the treating physician on patients' satisfaction with the patient-physician interaction. SCA usage in the ED had no significant effect on the primary endpoint (patients' satisfaction with the patient-physician interaction) or the prespecified secondary endpoints (patients' anxiety, patients' satisfaction with care, and physicians' satisfaction with the patient-physician interaction). Neither patients' nor physicians' satisfaction with their interaction increased with patients' prior use of the SCA, nor did patients' satisfaction with the care they received or physicians' satisfaction with the care they provided. Similarly, measures related to care efficiency, such as physicians' assessments of time to diagnosis and length of stay, showed no benefit of the SCA, with no differences between the treatment groups. This contrasts with the perceived effects of the SCA, as about one-third of patients in the intervention group reported that it had a positive impact on their patient-physician interaction and the care they received.

This discrepancy between the measured effects of the SCA and patients' perceptions aligns with previous literature on patients seeking health information online before receiving urgent care services. A survey-based observational study suggested a positive impact, with 150 out of 196 (76.5%) patients who had searched for information on their health problems before visiting an ED reporting that it improved the patient-physician relationship [5]. Meanwhile, an RCT primarily investigating the effect of web-based searches on the accuracy of patient-generated differential diagnoses found no evidence of an effect on its secondary endpoints, including patient-reported satisfaction with care, patients' and physicians' satisfaction with the patient-physician relationship, and patient anxiety [4].

We observed the same discrepancy between the perceived and measurable influence of the SCA among treating physicians. While physicians in the intervention group often considered their patients' use of the SCA helpful for certain tasks, such as diagnosis, this did not translate into a meaningful difference in their appraisal of the adequacy of time to diagnosis between the intervention and control groups.

These discrepancies raise further questions about why the perceived positive impact did not materialize in our study. Possible explanations include limitations in our choice of endpoints, effects that emerge only in specific subgroups, or differences in usage scenarios.

Our trial does not provide evidence of benefits associated with SCAs in acute care. However, it also does not indicate any negative effects, which are often the primary concern among providers and patients. More patients reported a significant decrease rather than an increase in anxiety after using the SCA, and only a few physicians noted instances of patients questioning their authority, with no difference between the groups.

As many as 2 in 5 physicians did not review the SCA report before seeing their patient, which may have diminished its impact on the patient-physician interaction. Allowing patients to bring up the report themselves and leaving it to the physician's discretion to consult a decision support or documentation tool more accurately reflects the reality of acute care. Therefore, we deliberately chose not to recommend that physicians engage with the SCA report.

This study has limitations. SCAs are used in various scenarios by users with diverse expectations [1,7,17,49], and our findings may not be generalizable to all these contexts. In our trial, patients used the SCA while in an ED, whereas its effect on anxiety levels might differ if used at home or outside a health care setting. Additionally, our study included only one of many available SCAs, and different SCAs may influence patients, physicians, and clinical care in distinct ways. We deliberately focused on investigating the SCA's effects on the

patient-physician relationship and satisfaction with care. Our study does not evaluate the investigated SCA's utility in guiding patients safely and efficiently through the health care system, as we recruited patients who had already decided to seek care at 1 of the 3 trial sites. Consequently, the participating patients' assessment of the SCA's utility does not reflect its potential value in the context of consulting the SCA before seeking care.

Furthermore, our findings may not capture important differences between subgroups. The positive or negative effects of SCAs might be specific to certain users, such as those with prior experience with online health information or individuals with hypochondriac traits. Notably, most patients in the intervention group had never used an SCA before. This could be a strength of our trial, as it allows us to estimate the effects of SCAs when used by a broader and more diverse population beyond early adopters. The impact of SCA use may differ between current users who independently choose to use these apps and those who are prompted to do so by health care providers. However, the proportion of trial participants with prior experience using SCAs was too small in our sample to allow for meaningful analysis. Additionally, all trial sites were located in highly urban areas, and the study participants were younger on average than the overall patient population presenting at these sites during the recruitment period. As a result, our patient sample is not fully representative of the broader acute care patient population.

Because of German employee data protection regulations, we were not permitted to match participating patients with their treating physicians. As a result, we are unable to assess whether physician-related variables may have confounded our results.

Postrecruitment, we observed a baseline age difference, with control group participants being, on average, 5 years older than those in the intervention group. Despite rigorous efforts, we found no violations of the recruitment procedure that could explain this discrepancy. As younger age correlated with higher usability ratings in our study, the younger average age of the intervention group may have biased our results in favor of the SCA.

While research continues to explore differences between search engine–based access to online health information and SCAs, advances in generative artificial intelligence have introduced new tools for both laypersons and professionals. Although some SCAs, including the one used in this trial, have demonstrated high usability, accuracy, and safety, our findings suggest that these attributes alone may not translate into added value for clinical care. This perspective is supported by a recent study that found no improvement in diagnostic quality when using a commercially available, physician-facing computerized diagnostic decision support system [74]. Even newer generative artificial intelligence–based apps designed to support clinical care may encounter similar challenges [75]. Possible avenues for advancement include integrating these intelligent support tools more seamlessly into existing processes or expanding their functionalities [76].

Conclusions

In summary, to our knowledge, we conducted the first non-industry–funded randomized controlled trial investigating the clinically relevant effects of using an SCA (Ada Health) in an acute care setting. Our trial provides no evidence of meaningful positive or negative effects of SCA use before the physician encounter on the patient-physician relationship or satisfaction with care. However, both patients and physicians more often perceived the SCA's influence as positive rather than negative. Thus, it remains an open question whether the perceived positive effects are unsubstantiated, were not captured by our chosen endpoints, or emerge only in specific subgroups or different usage scenarios. Our study did not identify negative effects of SCA use commonly described in the literature, such as inducing anxiety or eroding patients' trust in health care professionals. Our trial highlights the need for clinical research on mobile health apps, as high usability and reported accuracy did not necessarily translate into improved patient- and physician-reported outcomes.

Data Availability

Because of privacy concerns, we are unable to share the data publicly at the time of submission.

Authors' Contributions

MLS, MK, HN, and AS designed the study and supervised the project. MBA, MB, DK, CS, CW, and LS collected and double-checked the clinical data. MLS, MK, and LS loaded the data into the database. AT contributed to the literature search and the development of the trial's questionnaires. MLS, MK, LS, AS, HN, and SKP analyzed and interpreted the data. MM, FB, CH, MB, and KS advised on the study design and revised the manuscript. All authors vouched for the respective data and analysis, revised, approved the final version, and agreed to publish the manuscript. All authors had full access to all the data in the study and had final responsibility for the decision to submit for publication. The corresponding author attests that all listed authors meet authorship criteria and that no others meeting the criteria have been omitted.

Conflicts of Interest

HN invests in exchange-traded funds that hold shares of various companies in the technology and health care sectors, which do not gain financially from the publication of this manuscript but might do so in the future; is a member of the Digital Health

Multimedia Appendix 1

Additional analysis.
[DOCX File , 51 KB-Multimedia Appendix 1]

Multimedia Appendix 2

CONSORT 2010 checklist.
[PDF File (Adobe PDF File), 81 KB-Multimedia Appendix 2]

References

1. One in two EU citizens look for health information online. European Commission. Luxembourg City, Luxembourg. European Commission/Eurostat; Apr 6, 2021. URL: https://ec.europa.eu/eurostat/web/products-eurostat-news/-/edn-20210406-1 [accessed 2024-03-12]

2. Finney Rutten LJ, Blake KD, Greenberg-Worisek AJ, Allen SV, Moser RP, Hesse BW. Online health information seeking among us adults: measuring progress toward a healthy people 2020 objective. Public Health Rep. 2019;134(6):617-625. [FREE Full text] [doi: 10.1177/0033354919874074] [Medline: 31513756]

3. Baumann E, Czerwinski F, Rosset M, Seelig M, Suhr R. [How do people in Germany seek health information? Insights from the first wave of HINTS Germany]. Bundesgesundheitsblatt Gesundheitsforschung Gesundheitsschutz. Sep 2020;63(9):1151-1160. [doi: 10.1007/s00103-020-03192-x] [Medline: 32666180]

4. Martin SS, Quaye E, Schultz S, Fashanu OE, Wang J, Saheed MO, et al. A randomized controlled trial of online symptom searching to inform patient generated differential diagnoses. NPJ Digit Med. 2019;2:110. [FREE Full text] [doi: 10.1038/s41746-019-0183-0] [Medline: 31728417]

5. Cocco AM, Zordan R, Taylor DM, Weiland TJ, Dilley SJ, Kant J, et al. Dr Google in the ED: searching for online health information by adult emergency department patients. Med J Aust. Oct 15, 2018;209(8):342-347. [doi: 10.5694/mja17.00889] [Medline: 30107763]

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 Surveill. Jan 06, 2021;7(1):e22637. [doi: 10.2196/22637] [Medline: 33404515]

7.   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. Jan 30, 2020;22(1):e14679. [FREE Full text] [doi: 10.2196/14679] [Medline: 32012052]

8.   Kao C-K, Liebovitz DM. Consumer mobile health apps: current state, barriers, and future directions. PM R. May 2017;9(5S):S106-S115. [doi: 10.1016/j.pmrj.2017.02.018] [Medline: 28527495]

9.   Graafen MA, Sennekamp M, Messemaker A. „Also, im internet steht …" – wenn arztinnen und arzte auf internetinformierte patienten treffen. Zeitschrift für Allgemeinmedizin. May 1, 2021;97(5):210-214. [FREE Full text] [doi: 10.3238/zfa.2021.0210-0214]

10. Holzinger F, Oslislo S, Möckel M, Schenk L, Pigorsch M, Heintze C. Self-referred walk-in patients in the emergency department - who and why? Consultation determinants in a multicenter study of respiratory patients in Berlin, Germany. BMC Health Serv Res. Sep 10, 2020;20(1):848. [doi: 10.1186/s12913-020-05689-2] [Medline: 32912185]

11. Preiser C, Radionova N, Ög E, Koch R, Klemmt M, Müller R, et al. The doctors, their patients, and the symptom checker app: qualitative interview study with general practitioners in Germany. JMIR Hum Factors. Nov 18, 2024;11:e57360. [doi: 10.2196/57360] [Medline: 39556813]

12. Kujala S, Hörhammer I, Hänninen-Ervasti R, Heponiemi T. Health professionals' experiences of the benefits and challenges of online symptom checkers. Stud Health Technol Inform. Jun 16, 2020;270:966-970. [doi: 10.3233/SHTI200305] [Medline: 32570525]

13. Doherty-Torstrick ER, Walton KE, Fallon BA. Cyberchondria: parsing health anxiety from online behavior. Psychosomatics. 2016;57(4):390-400. [doi: 10.1016/j.psym.2016.02.002] [Medline: 27044514]

14. Van Riel N, Auwerx K, Debbaut P, Van Hees S, Schoenmakers B. The effect of Dr Google on doctor-patient encounters in primary care: a quantitative, observational, cross-sectional study. BJGP Open. May 17, 2017;1(2):bjgpopen17X100833. [FREE Full text] [doi: 10.3399/bjgpopen17X100833] [Medline: 30564661]

15. Using technology to ease the burden on primary care. Healthwatch England. Jan 22, 2019. URL: https://nds.healthwatch.co.uk/reports-library/using-technology-ease-burden-primary-care [accessed 2021-03-13]

16. Wetzel A-J, Koch R, Koch N, Klemmt M, Müller R, Preiser C, 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. [FREE Full text] [doi: 10.1177/20552076241231555] [Medline: 38434790]

17. Müller R, Klemmt M, Koch R, Ehni H-J, Henking T, Langmann E, et al. "That's just future medicine" - a qualitative study on users' experiences of symptom checker apps. BMC Med Ethics. Feb 16, 2024;25(1):17. [FREE Full text] [doi: 10.1186/s12910-024-01011-5] [Medline: 38365749]

18. 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. Jun 20, 2023;25:e46231. [doi: 10.2196/46231] [Medline: 37338970]

19. EPatient Analytics GmbH. EPatient Survey 2020. Health&Care Management. Berlin, Germany. Medizinisch Wissenschaftliche Verlagsgesellschaft mbH & Co. KG; Nov 3, 2020. URL: https://www.hcm-magazin.de/epatient-survey-2020-digital-health-studie-271773/ [accessed 2021-03-06]

20. Patienten-Navi. 116 117 - Der Patientenservice. Berlin, Germany. Kassenärztliche Bundesvereinigung KdöR URL: https://www.116117.de/de/patienten-navi.php [accessed 2025-03-24]

21. 111 Online. NHS England. NHS England URL: https://111.nhs.uk/ [accessed 2025-03-24]

22. Sutter Health teams up with Ada Health to improve patient care by delivering on-demand health care guidance. Vitals. Feb 11, 2019. URL: https://tinyurl.com/y3zyfcen [accessed 2021-11-13]

23. Check your symptoms. Kaiser Permanente Health Plan, Inc. URL: https://healthy.kaiserpermanente.org/health-wellness/health-encyclopedia/he.check-your-symptoms.hwsxchk [accessed 2021-11-13]

24. Symptom checker in kooperation mit infermedica. Sana Kliniken AG. Munich, Germany. Sana Kliniken AG URL: https://www.sana.de/unternehmen/digital/digitale-loesungen-im-einsatz/symptom-checker [accessed 2025-03-24]

25. KI-spezialist infermedica expandiert in Deutschland. Verband der Privaten Krankenversicherung e.V. Cologne, Germany. Verband der Privaten Krankenversicherung e.V; Feb 14, 2024. URL: https://www.pkv.de/verband/presse/meldungen/ki-spezialist-infermedica-gelingt-eintritt-in-den-gkv-markt/ [accessed 2025-03-24]

26. 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. Jun 2020;212(11):514-519. [doi: 10.5694/mja2.50600] [Medline: 32391611]

27. Semigran HL, Linder J, Gidengil C, Mehrotra A. Evaluation of symptom checkers for self diagnosis and triage: audit study. BMJ. Jul 08, 2015;351:h3480. [FREE Full text] [doi: 10.1136/bmj.h3480] [Medline: 26157077]

28. 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]

XSL•FO

RenderX

29. Hennemann S, Kuhn S, Witthöft M, Jungmann SM. Diagnostic performance of an app-based symptom checker in mental disorders: comparative study in psychotherapy outpatients. JMIR Ment Health. Jan 31, 2022;9(1):e32832. [doi: 10.2196/32832] [Medline: 35099395]

30. Knitza J, Hasanaj R, Beyer J, Ganzer F, Slagman A, Bolanaki M, et al. Comparison of two symptom checkers (Ada and Symptoma) in the emergency department: randomized, crossover, head-to-head, double-blinded study. J Med Internet Res. Aug 20, 2024;26:e56514. [FREE Full text] [doi: 10.2196/56514] [Medline: 39163594]

31. Knitza J, Tascilar K, Fuchs F, Mohn J, Kuhn S, Bohr D, et al. Diagnostic accuracy of a mobile AI-based symptom checker and a web-based self-referral tool in rheumatology: multicenter randomized controlled trial. J Med Internet Res. Jul 23, 2024;26:e55542. [doi: 10.2196/55542] [Medline: 39042425]

32. Fraser H, Crossland D, Bacher I, Ranney M, Madsen T, Hilliard R. Comparison of diagnostic and triage accuracy of Ada Health and WebMD symptom checkers, ChatGPT, and physicians for patients in an emergency department: clinical data analysis study. JMIR Mhealth Uhealth. Oct 03, 2023;11:e49995. [doi: 10.2196/49995] [Medline: 37788063]

33. 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. May 10, 2022;24(5):e31810. [doi: 10.2196/31810] [Medline: 35536633]

34. Riboli-Sasco E, El-Osta A, Alaa A, Webber I, Karki M, El Asmar ML, et al. Triage and diagnostic accuracy of online symptom checkers: systematic review. J Med Internet Res. Jun 02, 2023;25:e43803. [doi: 10.2196/43803] [Medline: 37266983]

35. Meer A, Rahm P, Schwendinger M, Vock M, Grunder B, Demurtas J, et al. A symptom-checker for adult patients visiting an interdisciplinary emergency care center and the safety of patient self-triage: real-life prospective evaluation. J Med Internet Res. Jun 27, 2024;26:e58157. [doi: 10.2196/58157] [Medline: 38809606]

36. Hammoud M, Douglas S, Darmach M, Alawneh S, Sanyal S, Kanbour Y. Evaluating the diagnostic performance of symptom checkers: clinical vignette study. JMIR AI. Apr 29, 2024;3:e46875. [doi: 10.2196/46875] [Medline: 38875676]

37. Liu V, Kaila M, Koskela T. Triage accuracy and the safety of user-initiated symptom assessment with an electronic symptom checker in a real-life setting: instrument validation study. JMIR Hum Factors. Sep 26, 2024;11:e55099. [FREE Full text] [doi: 10.2196/55099] [Medline: 39326038]

38. 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. [FREE Full text] [doi: 10.1371/journal.pone.0199284] [Medline: 29944708]

39. Winn AN, Somai M, Fergestrom N, Crotty BH. Association of use of online symptom checkers with patients' plans for seeking care. JAMA Netw Open. Dec 02, 2019;2(12):e1918561. [FREE Full text] [doi: 10.1001/jamanetworkopen.2019.18561] [Medline: 31880791]

40. Chambers D, Cantrell A, Johnson M, Preston L, Baxter SK, Booth A, et al. Digital and Online Symptom Checkers and Assessment Services for Urgent Care to Inform a New Digital Platform: A Systematic Review. Southampton, UK. NIHR Journals Library; Aug 2019.

41. Gellert GA, Orzechowski PM, Price T, Kabat-Karabon A, Jaszczak J, Marcjasz N, et al. A multinational survey of patient utilization of and value conveyed through virtual symptom triage and healthcare referral. Front Public Health. 2022;10:1047291. [FREE Full text] [doi: 10.3389/fpubh.2022.1047291] [Medline: 36817183]

42. Gellert GA, Kabat-Karabon A, Gellert GL, Rasławska-Socha J, Gorski S, Price T, et al. The potential of virtual triage AI to improve early detection, care acuity alignment, and emergent care referral of life-threatening conditions. Front Public Health. 2024;12:1362246. [FREE Full text] [doi: 10.3389/fpubh.2024.1362246] [Medline: 38807993]

43. Ronicke S, Hirsch MC, Türk E, Larionov K, Tientcheu D, Wagner AD. Can a decision support system accelerate rare disease diagnosis? Evaluating the potential impact of Ada DX in a retrospective study. Orphanet J Rare Dis. Mar 21, 2019;14(1):69. [FREE Full text] [doi: 10.1186/s13023-019-1040-6] [Medline: 30898118]

44. 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. Sep 19, 2022;10(9):e38364. [doi: 10.2196/38364] [Medline: 36121688]

45. Gottliebsen K, Petersson G. Limited evidence of benefits of patient operated intelligent primary care triage tools: findings of a literature review. BMJ Health Care Inform. May 2020;27(1):e100114. [doi: 10.1136/bmjhci-2019-100114] [Medline: 32385041]

46. 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. [doi: 10.3389/fmed.2022.1040926] [Medline: 36687416]

47. 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. Aug 17, 2022;5(1):118. [doi: 10.1038/s41746-022-00667-w] [Medline: 35977992]

48. Radionova N, Ög E, Wetzel A-J, Rieger MA, Preiser C. Impacts of symptom checkers for laypersons' self-diagnosis on physicians in primary care: scoping review. J Med Internet Res. May 29, 2023;25:e39219. [doi: 10.2196/39219] [Medline: 37247214]

49.   Wetzel A-J, Preiser C, Müller R, Joos S, Koch R, Henking T, et al. Unveiling usage patterns and explaining usage of symptom checker apps: explorative longitudinal mixed methods study. J Med Internet Res. Dec 09, 2024;26:e55161. [doi: 10.2196/55161] [Medline: 39652845]

50.   Snow G. Blockrand: randomization for block random clinical trials. The Comprehensive R Archive Network. Apr 6, 2020. URL: https://cran.r-project.org/web/packages/blockrand/index.html [accessed 2025-03-24]

51.   Franke T, Attig C, Wessel D. A personal resource for technology interaction: development and validation of the Affinity for Technology Interaction (ATI) scale. International Journal of Human–Computer Interaction. Mar 30, 2018;35(6):456-467. [doi: 10.1080/10447318.2018.1456150]

52.   European Commission. European Health Interview Survey (EHIS Wave 2): Methodological Manual (2013 Edition). Luxembourg City, Luxembourg. Publications Office of the European Union; Jul 18, 2013:13-14.

53.   Beierlein C, Kovaleva A, Kemper CJ, Rammstedt B. Allgemeine Selbstwirksamkeit Kurzskala (ASKU). Open Test Archive. 2014. URL: https://www.testarchiv.eu/de/test/9006490 [accessed 2025-03-24]

54.   Soellner R, Huber S, Reder M. The concept of eHealth literacy and its measurement: German translation of the eHEALS. Journal of Media Psychology. 2014;26:29-38. [FREE Full text] [doi: 10.1027/1864-1105/a000104]

55.   Ada Health. URL: https://ada.com/ [accessed 2025-03-24]

56.   Knitza J, Muehlensiepen F, Ignatyev Y, Fuchs F, Mohn J, Simon D, 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] [Medline: 35273944]

57.   Mehl A, Bergey F, Cawley C, Gilsdorf A. Syndromic surveillance insights from a symptom assessment app before and during COVID-19 measures in Germany and the United Kingdom: results from repeated cross-sectional analyses. JMIR Mhealth Uhealth. Oct 09, 2020;8(10):e21364. [FREE Full text] [doi: 10.2196/21364] [Medline: 32997640]

58.   Ćirković A. Evaluation of four artificial intelligence-assisted self-diagnosis apps on three diagnoses: two-year follow-up study. J Med Internet Res. Dec 04, 2020;22(12):e18097. [FREE Full text] [doi: 10.2196/18097] [Medline: 33275113]

59.   Cotte F, Mueller T, Gilbert S, Blümke B, Multmeier J, Hirsch MC, et al. Safety of triage self-assessment using a symptom assessment app for walk-in patients in the emergency care setting: observational prospective cross-sectional study. JMIR Mhealth Uhealth. Mar 28, 2022;10(3):e32340. [doi: 10.2196/32340] [Medline: 35343909]

60.   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. Nov 30, 2020;22(11):e20549. [FREE Full text] [doi: 10.2196/20549] [Medline: 33170799]

61.   Gilbert S, Fenech M, Upadhyay S, Wicks P, Novorol C. Quality of condition suggestions and urgency advice provided by the Ada symptom assessment app evaluated with vignettes optimised for Australia. Aust J Prim Health. Oct 2021;27(5):377-381. [doi: 10.1071/PY21032] [Medline: 34706813]

62.   Blanchard CG, Ruckdeschel JC, Fletcher BA, Blanchard EB. The impact of oncologists' behaviors on patient satisfaction with morning rounds. Cancer. Jul 15, 1986;58(2):387-393. [doi: 10.1002/1097-0142(19860715)58:2<387::aid-cncr2820580233>3.0.co;2-3] [Medline: 3719533]

63.   Zandbelt LC, Smets EMA, Oort FJ, Godfried MH, de Haes HCJM. Satisfaction with the outpatient encounter: a comparison of patients' and physicians' views. J Gen Intern Med. Nov 2004;19(11):1088-1095. [FREE Full text] [doi: 10.1111/j.1525-1497.2004.30420.x] [Medline: 15566437]

64.   Schmidt J, Lamprecht F, Wittmann WW. [Satisfaction with inpatient management. Development of a questionnaire and initial validity studies]. Psychother Psychosom Med Psychol. Jul 1989;39(7):248-255. [Medline: 2762479]

65.   Attkisson CC, Greenfield TK. Client Satisfaction Questionnaire-8 and Service Satisfaction Scale-30. In: The Use of Psychological Testing for Treatment Planning and Outcome Assessment. Hillsdale, NJ. Lawrence Erlbaum Associates, Inc; 1994:402-420.

66.   Abend R, Dan O, Maoz K, Raz S, Bar-Haim Y. Reliability, validity and sensitivity of a computerized visual analog scale measuring state anxiety. J Behav Ther Exp Psychiatry. Dec 2014;45(4):447-453. [doi: 10.1016/j.jbtep.2014.06.004] [Medline: 24978117]

67.   R Core Team. R: a language and environment for statistical computing. R Foundation for Statistical Computing. Vienna, Austria. R Foundation for Statistical Computing; 2024. URL: https://www.R-project.org/ [accessed 2025-03-24]

68.   Wickham H, Averick M, Bryan J, Chang W, McGowan L, François R, et al. Welcome to the Tidyverse. JOSS. Nov 2019;4(43):1686. [FREE Full text] [doi: 10.21105/joss.01686]

69.   Bates D, Mächler M, Bolker B, Walker S. Fitting linear mixed-effects models using lme4. J Stat Soft. 2015;67(1):1-48. [FREE Full text] [doi: 10.18637/jss.v067.i01]

70.   Lüdecke D, Ben-Shachar M, Patil I, Makowski D. Extracting, computing and exploring the parameters of statistical models using R. JOSS. Sep 2020;5(53):2445. [FREE Full text] [doi: 10.21105/joss.02445]

71.   van Buuren S, Groothuis-Oudshoorn K. mice: multivariate imputation by chained equations in R. J Stat Soft. 2011;45(3):1-67. [FREE Full text] [doi: 10.18637/jss.v045.i03]

72.   Napierala H, Kopka M, Altendorf MB, Bolanaki M, Schmidt K, Piper SK, et al. Examining the impact of a symptom assessment application on patient-physician interaction among self-referred walk-in patients in the emergency department

(AKUSYM): study protocol for a multi-center, randomized controlled, parallel-group superiority trial. Trials. Sep 20, 2022;23(1):791. [doi: 10.1186/s13063-022-06688-w] [Medline: 36127742]

73. Schulz KF, Altman DG, Moher D, CONSORT Group. CONSORT 2010 statement: updated guidelines for reporting parallel group randomised trials. BMJ. Mar 23, 2010;340:c332. [doi: 10.1136/bmj.c332] [Medline: 20332509]

74. Hautz WE, Marcin T, Hautz SC, Schauber SK, Krummrey G, Müller M, et al. Diagnoses supported by a computerised diagnostic decision support system versus conventional diagnoses in emergency patients (DDX-BRO): a multicentre, multiple-period, double-blind, cluster-randomised, crossover superiority trial. Lancet Digit Health. Feb 2025;7(2):e136-e144. [doi: 10.1016/S2589-7500(24)00250-4] [Medline: 39890244]

75. Goh E, Gallo R, Hom J, Strong E, Weng Y, Kerman H, et al. Large language model influence on diagnostic reasoning: a randomized clinical trial. JAMA Netw Open. Oct 01, 2024;7(10):e2440969. [FREE Full text] [doi: 10.1001/jamanetworkopen.2024.40969] [Medline: 39466245]

76. Kostopoulou O, Delaney B. AI for medical diagnosis: does a single negative trial mean it is ineffective? Lancet Digit Health. Feb 2025;7(2):e108-e109. [doi: 10.1016/j.landig.2025.01.005] [Medline: 39890240]

Abbreviations

CCM: Charité – Universitätsmedizin Berlin, Campus Mitte in Berlin-Mitte
CONSORT: Consolidated Standards of Reporting Trials
CSQ-8: 8-item Client Satisfaction Questionnaire
CVK: Campus Virchow-Klinikum in Berlin-Wedding
ED: emergency department
JKB: Jüdisches Krankenhaus Berlin in Berlin-Gesundbrunnen
MTS: Manchester Triage System
PSQ: Patient Satisfaction Questionnaire
SCA: symptom checker app
ZUF-8: 8-item Fragebogen zur Patientenzufriedenheit