

Impact of a symptom checker application on patient-physician interaction among self-referred walk-in patients in the emergency department: a multi-centre, parallel-group, randomized, controlled trial

Malte L. Schmieding, Marvin Kopka, Myrto Bolanaki, Hendrik Napierala, Maria B. Altendorf, Doreen Kuschick, Sophie K. Piper, Lennart Scatturin, Konrad Schmidt, Claudia Schorr, Alica Thissen, Cornelia Wäscher, Christoph Heintze, Martin Möckel, Felix Balzer, Anna Slagman

Submitted to: Journal of Medical Internet Research
on: August 11, 2024

Disclaimer: © The authors. All rights reserved. This is a privileged document currently under peer-review/community review. Authors have provided JMIR Publications with an exclusive license to publish this preprint on its website for review purposes only. While the final peer-reviewed paper may be licensed under a CC BY license on publication, at this stage authors and publisher expressly prohibit redistribution of this draft paper other than for review purposes.

Table of Contents

Original Manuscript..... 5

Supplementary Files..... 33

 Figures 34

 Figure 1..... 35

 Multimedia Appendixes 36

 Multimedia Appendix 1..... 37

 CONSORT (or other) checklists..... 38

 CONSORT (or other) checklist 0..... 38

Impact of a symptom checker application on patient-physician interaction among self-referred walk-in patients in the emergency department: a multi-centre, parallel-group, randomized, controlled trial

Malte L. Schmieding¹; Marvin Kopka^{1, 2}; Myrto Bolanaki³; Hendrik Napierala⁴; Maria B. Altendorf³; Doreen Kuschick⁴; Sophie K. Piper^{1, 5, 2}; Lennart Scatturin¹; Konrad Schmidt⁴; Claudia Schorr⁴; Alica Thissen¹; Cornelia Wäscher⁴; Christoph Heintze⁴; Martin Möckel³; Felix Balzer¹; Anna Slagman³

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

²Emergency and Acute Medicine and Health Services Research in Emergency Medicine (CVK, CCM) Charité – Universitätsmedizin Berlin, corporate member of Freie Universität Berlin and Humboldt-Universität zu Berlin Berlin DE

³Institute of General Practice and Family Medicine Charité – Universitätsmedizin Berlin, corporate member of Freie Universität Berlin and Humboldt-Universität zu Berlin Berlin DE

⁴Berlin Institute of Health at Charité – Universitätsmedizin Berlin Berlin DE

Corresponding Author:

Malte L. Schmieding

Institute of Medical Informatics

Charité – Universitätsmedizin Berlin, corporate member of Freie Universität Berlin and Humboldt-Universität zu Berlin Berlin

DE

Abstract

Background: Symptom-checker applications (SCAs) are layperson-facing tools providing advice on whether and where to seek care (triage) and / or possible diagnoses. Previous research focused on assessing accuracy and safety of triage advice and diagnostic suggestions of the myriad of available SCAs, and their usability.

Objective: As studies on SCAs' impact remain scarce this trial evaluates an SCA's effects on patient-physician interaction and satisfaction with acute care.

Methods: This parallel-group, randomised, controlled trial was carried out at two emergency departments (EDs) of an academic medical centre and an emergency practice in Berlin, Germany. Low-acuity patients seeking care in these three trial sites were randomly assigned to either self-assess their health complaints with a widely available SCA prior to their first encounter with the treating physician or to care as usual. Our primary endpoint was patients' satisfaction with the patient-physician interaction as measured by the Patient Satisfaction Questionnaire (PSQ). Secondary objectives included patients' anxiety and their satisfaction with care.

Results: A total of 363 patients were included (intervention: 173, control: 190) in the modified intention-to-treat analysis. PSQ in the intervention group was similar to the control group (mean (SD): M=78.5 (20.0) vs. 80.8 (19.6); estimated difference: -2.4, 95%-CI: -6.3 to 1.1, p = 0.24). Secondary outcomes revealed no significant group differences, either.

Conclusions: Our study provides no evidence for SCAs improving the patient-physician interaction or satisfaction with care in acute care. Clinical Trial: The trial was registered in the German Clinical Trial Registry (DRKS00028598).

(JMIR Preprints 11/08/2024:64028)

DOI: <https://doi.org/10.2196/preprints.64028>

Preprint Settings

1) Would you like to publish your submitted manuscript as preprint?

Please make my preprint PDF available to anyone at any time (recommended).

Please make my preprint PDF available only to logged-in users; I understand that my title and abstract will remain visible to all users.

Only make the preprint title and abstract visible.

✓ **No, I do not wish to publish my submitted manuscript as a preprint.**

2) If accepted for publication in a JMIR journal, would you like the PDF to be visible to the public?

✓ **Yes, please make my accepted manuscript PDF available to anyone at any time (Recommended).**

Yes, but please make my accepted manuscript PDF available only to logged-in users; I understand that the title and abstract will remain visible to the public.

Yes, but only make the title and abstract visible (see Important note, above). I understand that if I later pay to participate in a JMIR journal, my accepted manuscript PDF will be made available to the public.

No. Please do not make my accepted manuscript PDF available to anyone.



Original Manuscript

Impact of a symptom checker application on patient-physician interaction among self-referred walk-in patients in the emergency department: a multi-centre, parallel-group, randomized, controlled trial

Authors: Malte L. Schmieding¹, Marvin Kopka^{1,2}, Myrto Bolanaki³, Hendrik Napierala⁴, Maria B. Altendorf³, Doreen Kuschick⁴, Sophie K. Piper^{1,5,6}, Lennart Scatturin¹, Konrad Schmidt⁴, Claudia Schorr⁴, Alica Thissen¹, Cornelia Wäscher⁴, Christoph Heintze⁴, Martin Möckel³, Felix Balzer¹, Anna Slagman³

Affiliations:

1: Charité – Universitätsmedizin Berlin, corporate member of Freie Universität Berlin and Humboldt-Universität zu Berlin, Institute of Medical Informatics, Charitéplatz 1, 10117 Berlin, Germany

2: Division of Ergonomics, Department of Psychology and Ergonomics (IPA), Technische Universität Berlin, Straße des 17. Juni 135, 10623 Berlin, Germany

3: Charité – Universitätsmedizin Berlin, corporate member of Freie Universität Berlin and Humboldt-Universität zu Berlin, Emergency and Acute Medicine and Health Services Research in Emergency Medicine (CVK, CCM), Charitéplatz 1, 10117 Berlin, Germany

4: Charité – Universitätsmedizin Berlin, corporate member of Freie Universität Berlin and Humboldt-Universität zu Berlin, Institute of General Practice and Family Medicine, Charitéplatz 1, 10117 Berlin, Germany

5: Charité – Universitätsmedizin Berlin, corporate member of Freie Universität Berlin and Humboldt-Universität zu Berlin, Institute of Biometry and Clinical Epidemiology, Charitéplatz 1, 10117, Berlin, Germany

6: Berlin Institute of Health at Charité – Universitätsmedizin Berlin, Charitéplatz 1, 10117, Berlin, Germany

Corresponding author:

Malte L. Schmieding, malte.schmieding@charite.de

Keywords: digital health; triage; symptom checker; patient-centered care; eHealth apps; mobile phone; decision support systems; consumer health information; health literacy; randomised controlled trials; null results; emergency care; patient-physician-interaction; patient satisfaction



Abstract

Background

Symptom-checker applications (SCAs) are layperson-facing tools providing advice on whether and where to seek care and / or possible diagnoses. Previous research focused on assessing accuracy and safety of their advice and suggestions in available SCAs, and their usability. Studies on SCAs' impact on clinical care, like the patient-physician interaction and satisfaction with care remain scarce.

Objective

To evaluate one SCA's effects on satisfaction with the patient-physician interaction in acute care settings. Additionally, we examined the influence on patients' anxiety and trust in the treating physician.

Methods

This parallel-group, randomised, controlled trial was carried out at two emergency departments (EDs) of an academic medical centre and an emergency practice in Berlin, Germany. Low-acuity patients seeking care in these trial sites were randomly assigned to self-assess their health complaints with a widely available, commercial SCA (Ada Health) prior to their first encounter with the treating physician or to care as usual. Our primary endpoint was patients' satisfaction with the patient-physician interaction as measured by the Patient Satisfaction Questionnaire (PSQ). Secondary outcomes investigated patients' satisfaction with care, their anxiety, and their physicians' satisfaction with the patient-physician interaction. We used linear mixed models to determine the statistical significance of primary and secondary outcomes. Exploratory, descriptive analyses included patients' and physicians' perceptions of the SCA's utility, and the rate of patients questioning their physician's authority.

Results

Between April 11th 2022 and January 25th 2023 we approached 665 patients. 363 patients were included (intervention: 173, control: 190) in the intention-to-treat analysis of the primary outcome. PSQ in the intervention group was similar to the control group (mean (SD): M=78.5 (20.0) vs. 80.8 (19.6); estimated difference: -2.4, 95%-CI: -6.3 to 1.1, p=0.24). Secondary outcomes (patients' and physicians' satisfaction with care, patient anxiety) revealed no significant group differences, either. Patients in the intervention group more often indicated that the SCA had a beneficial (40.2%, 66/164) than detrimental (1.8%, 3/164) impact on the patient-physician-interaction, with most reporting no effect (57.9%, 95/164). We observed similar numbers regarding the SCA's perceived effect on care. In both groups, physicians rarely reported their authority had been questioned by a patient (intervention: 1.1% (2/188); control: 2.2% (4/184)). Physicians more often indicated the SCA had been helpful than unhelpful, but the absolute majority indicated it had been neither helpful nor unhelpful for the encounter.

Conclusions

We found no evidence for the SCA to improve the satisfaction with the patient-physician interaction and care in acute care setting. In contrast, patients and their treating physicians described the SCA's impact as mostly beneficial. Our study did not reveal negative effects of SCA use commonly described in the literature such as inducing anxiety or eroding patients' trust in healthcare professionals.

Trial Registration

The trial was registered in the German Clinical Trial Registry (DRKS00028598).

Funding

German Ministry of Health





Introduction

It has become common in the general population to seek health-related information online. In the European Union more than one in two citizens have searched for health-related information online in the last three months prior to being surveyed.¹ Similarly, 74.4% of US adults consulted the internet first for obtaining health information during their most recent inquiry.² In a German panel study a fifth of the respondents named the internet as their primary source of health information.³ In particular, seeking health-related information online is common prior to consulting medical services among low-urgency acute care patients.^{4,5} While qualitative studies outline concerns from both patients⁶⁻⁸ and healthcare professionals⁹⁻¹² regarding online information exacerbating patient anxiety¹³ and undermining the patient-physician-relationship, two 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 are symptom checker applications (SCAs). These consumer apps offer suggestions on potential diagnoses and/or an urgency assessment for the self-reported signs and symptoms the users enter. SCAs face similar concerns to web-based health information seeking behavior in general as described above, especially with regards to inducing anxiety¹⁵⁻¹⁷. Studies estimate the proportion of SCA users in the German population between 6 and 13%.^{16,18,19} Some national healthcare services,^{20,21} healthcare systems,^{22,23} hospital networks²⁴ and insurance companies²⁵ have already integrated SCAs into their service pathways.

So far, research studies have commonly focused on assessing the accuracy and safety of SCA advice²⁶⁻³⁸, their (potential) impact on patient journeys and allocation^{28,39-43}, and users' experience and expectations^{6,7,17,30,38,41,44} but not the SCAs' actual impact on patient-centred, clinically relevant outcomes. Four recent reviews concluded that more research on SCAs' impacts in real-life settings is needed to estimate their utility.⁴⁵⁻⁴⁸ To address this research gap, we conducted a multicentre, randomized controlled trial assessing 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 commonly feature in discussions about SCAs' and online health information's impact on care.^{4,5,9,10,49} We bracketed out questions regarding the SCA's utility for improving patient allocation. The primary hypothesis was that the intervention would improve patients' satisfaction with their interaction with the treating physician. Secondary hypotheses included patients' satisfaction with the health care received and their anxiety levels.

Methods

Study design, setting and participants

We conducted a multicentre, controlled parallel-group trial with balanced randomization, at three study sites in Berlin, Germany. Two study sites were 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 offer a wide spectrum of non-pediatric care for approximately 50,000 patient encounters yearly and operate an adjunct acute medical admissions ward. The third site was an emergency practice run by Berlin's association of statutory health insurance physicians (Kassenärztliche Vereinigung Berlin) adjunct to the ED of a local hospital (JKB: Jüdisches Krankenhaus Berlin in Berlin-Gesundbrunnen). The emergency practice is an outpatient clinic offering urgent care outside of the office hours of regular outpatient clinics to around 4,000 walk-in patients per year. It is commonly staffed by one specialist physician and one medical assistant. The emergency practice caters to self-referred patients not requiring care in the adjunct hospital-run emergency department, located in the same building. This stratification of emergency care is intended to ensure that inpatient resources are used only for urgently triaged patients. Key inclusion criteria were: self-referred walk-in patients aged 18 years and older with sufficient German or English language proficiency, the ability to provide informed consent, and an assigned a treatment urgency rating of yellow, green or blue according to the Manchester Triage System (MTS 3-5) by the triage nurse. Exclusion criteria were: patient treated without waiting time; cause of the chief complaint already known according to the patient; patient requiring isolation; patient unable to handle a tablet computer according to the patient's own or the study personnel's judgment; patient already consulted a SCA for their current complaints prior to seeking care.

Randomization and masking

Participants were randomly assigned to either using the SCA before the first encounter with a treating physician or care as usual (1:1 ratio). Balanced block randomization with variable block lengths of eight, ten and twelve was used. The recruiting study personnel were blinded to block size. Each trial site received its own allocation sequence in sequentially numbered, opaque and sealed envelopes. The allocation sequence was generated by the Institute of Medical Informatics by a researcher not actively involved in recruitment (MLS) using the R package blockrand⁵⁰. Allocation was concealed up to the point of randomization which was immediately after the patient consented to participating in the trial. After randomization participants, study personnel and health-care providers were not masked to group assignment due to the intervention design.

Procedures

Following initial triage, patients underwent eligibility screening conducted by the study personnel (study nurse, student research assistant or study physician) and, upon providing informed consent, were randomly assigned to either the control or intervention group (see details on randomization below). All participants completed a baseline survey (baseline anxiety, prior use of SCA, and affinity for technology interaction⁵¹) prior to the first encounter with the treating physician. A survey on socio-demographic and other variables not subject to change during a patient's stay (age, sex, native language, country of residence, level of education, net household income, frequency of using the internet, tablet computers and smartphones, self-perceived health and chronic morbidity⁵², self-efficacy⁵³ eHealth-Literacy⁵⁴) was handed out at a suitable time during their stay, ie either before or after the encounter with the treating physician (sociodemographic survey). Furthermore, after seeing the treating physician or in exceptional cases within 72 hours thereafter, all participants were asked to answer a post-encounter survey. Participants in the intervention group completed

the self-assessment of their symptoms with SCA after taking the baseline survey but prior to their first encounter with the treating physician. Following SCA use and preceding consultation with a treating physician, participants in the intervention group completed a post-SCA survey on their experience with the SCA and their level of anxiety. We administered all surveys through a tablet computer. Study personnel instructed the participants on using the tablet.

Study personnel ensured that a print-out of the SCA summary reports were available to the treating physicians at the time of their first encounter with patients in the intervention group. However, in order not to interfere with the patient-physician-relationship and care provision, study personnel neither encouraged nor discouraged the treating physicians to engage with this summary report. All participating physicians were aware that such summary reports will be in place for patients in the intervention group. After providing care to the patients, the participating physicians completed the physician-sided, paper-based survey (physician-sided PSQ, physician's satisfaction with care provided, whether the patient questioned their authority, physician's appraisal of the SCA's helpfulness and impact on the patient encounter).

At the trial site JKB all patients presenting during recruiting days were screened for eligibility. Due to greater patient volume, we were unable to screen all patients for eligibility at trial sites CCM and CVK.

The SCA used in the intervention (© 2024 Ada Health GmbH: www.ada.com) was developed independently of the researchers involved in this trial. We selected this particular commercial SCA out of the many available based on the available literature on its reported (diagnostic and triage) accuracy, the safety of its advice, its usability, breadth of conditions and chief complaints covered^{28,33,44,55}, and its availability in both German and English. It requires users to detail basic demographic information (such as age and sex), past medical history (such as smoking behavior and prior diagnoses) and information on the current medical complaints. In this first step, the user can select an unlimited number of symptoms⁵⁶. The user then provides further information by answering a sequence of closed questions with binary or multiple answer options which the SCA presents in a conversational manner. Such "conversations" last about six to eight minutes.^{29,55,57} Then, the SCA provides an assessment of the urgency of the complaints and suggests one to five probable causes (diagnostic suggestions), which are illustrated using a Sankey diagram. During the time we conducted our trial, the SCA inquired neither about the users' intent to seek care nor their own urgency assessment of their complaints, nor the diagnoses they suspected. The SCA's report also summarizes the findings affirmed or denied by the user. The SCA generates its suggestions using a medical knowledge base, that hard-codes libraries of signs and symptoms, and diagnoses and their relationships, and a Bayesian network algorithm.⁵⁸ The company behind Ada Health describes the algorithm as "artificial intelligence".⁵⁹ Further details of the SCA's underlying algorithm are not publicly available. Ada Health is the second most frequently used SCA in Germany, according to 2024 study¹⁶, and according to the developers also widely used in other countries, namely Australia⁶⁰.

In the care as usual group patients only used the tablet to fill out the questionnaires (baseline, socio-demographic, and post-encounter survey) and could not self-assess their symptoms with the SCA via the tablet. We did not prohibit patients in both groups from searching for online health information or for using a SCA with their own devices during the participation in the study.

The study personnel observed patients closely for any signs of discomfort during the study procedures. In both study groups clinical routine care was conducted as usual. Study procedures were interrupted for clinical

interventions or if otherwise necessary.

Outcomes

Primary Outcome

The primary outcome was participants' satisfaction with their interaction with the treating physician assessed by the Patient Satisfaction Questionnaire (PSQ)^{61,62}, an instrument consisting of visual analogue scales ranging from 0 to 100 for each of 5 items. A Participant's satisfaction is defined as the average of all five items, with higher values indicating greater satisfaction. We collected the participants' responses on the primary outcome in the post-encounter survey prior to discharge or, in exceptional cases, within 72 hours of discharge in case the participants left the trial site without completing the post-encounter questionnaire. We considered a mean difference of 5 points (Scale 0 - 100) between treatment groups as clinically relevant.

Secondary and exploratory outcomes

Participants' satisfaction with the care received was measured using the "Fragebogen zur Patientenzufriedenheit" (ZUF-8)⁶³ instrument, a German version of the Client Satisfaction Questionnaire (CSQ-8)⁶⁴. The ZUF-8's scale ranges from 8 to 32 points, with higher values indicating greater satisfaction. As some participants did not respond to all 8 items of the ZUF-8 (each having a scale of 1 to 4), we calculated the average value of items to which the participant responded. Thus, we report our results for the ZUF-8 on a scale from 1 to 4. This is a deviation from the protocol where we planned no such modification to account for missing values. We expected a difference of one point in care satisfaction as measured by the ZUF-8 original scale between the intervention and control, corresponding to 0.125 points on the scale ranging from 1 to 4.

Participants' anxiety was measured up to three times during the course of the trial: initially after recruitment (baseline), after using the SCA for participants in the intervention group, and a finally after the patient-physician-interaction. We used a visual analogue scale with a range from 0 to 100 with higher values indicating greater anxiety⁶⁵. Physicians' satisfaction with the patient-physician interaction was measured using the physician version of the PSQ^{61,62}, which rephrases the PSQ items to the physician perspective.

As further exploratory outcomes we report participants' perceived effect of the SCA on the patient-physician interaction and patient care (intervention group only; 2 5-point Likert scales), the helpfulness of the SCA's report according to physicians (intervention group only; 5 3-point Likert scales), and the physicians' satisfaction with the care they delivered (regarding overall satisfaction, time to diagnosis, and patient's length-of-stay; visual analogue scale ranging from 0 - 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 intervention and control group would be clinically relevant: In the literature we found a standard deviation of the patient-facing PSQ of 14 to 17 points.⁶² Assuming a standardized mean difference of 0.3 and equal variance, and considering a two-sided alpha of 0.05, a power of 0.80 and a medium-sized drop-out rate of 20%, a total of 440 patients were needed (i.e., 220 in each arm of the trial). For the purpose of sample size calculation, we conservatively employed a two-sample 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 given as mean with standard deviation (SD), median with the interquartile range (IQR) or frequency and proportion depending on scale and distribution.

We conducted our primary analysis according to the modified intention-to-treat of principle, including all randomly assigned patients who reported on at least one item related to the respective primary or secondary outcome measure in the post-encounter survey. We conducted subgroup analyses by study site as pre-planned.

We analyzed the primary outcome (patients' satisfaction with the patient-physician interaction as measured by the PSQ) by means of a linear mixed model with fixed effect for intervention and study site as random effect (including only participants who responded to at least one item of the PSQ). We report the group means and standard deviations, the estimator of the linear mixed model, 95% confidence intervals, and the respective p value. We conducted statistical analyses with *R*, version 4.4.0⁶⁶, using the tidyverse⁶⁷ for data cleaning. We bootstrapped confidence intervals and calculated p values using the *R* packages lme4⁶⁸ and parameters⁶⁹, respectively. For secondary outcomes (patients' satisfaction with the care received, change in anxiety levels after SCA use, proportion of participants more anxious after the physician encounter than at baseline, physicians' satisfaction with patient-physician interaction) our analyses follow 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 employed as sensitivity analysis (see Supplementary Table 3). Regarding exploratory outcomes we report descriptive statistics only.

The researchers statistically assessing the primary outcome were blinded to group assignment. Though 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 mice⁷⁰.

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 was 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 prior to conducting analyses, limiting access to all trial data, and storing and processing potentially re-identifiable 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 has been previously published.⁷¹ This manuscript follows the Consolidated Standards of Reporting Trials (CONSORT) statement.⁷² We compensated each participating patient with a 10€ gift voucher redeemable at over 300 brands and e-commerce platforms. Treating physicians received a 5€ gift voucher for every survey they completed on an encounter with a patient enrolled in the trial.

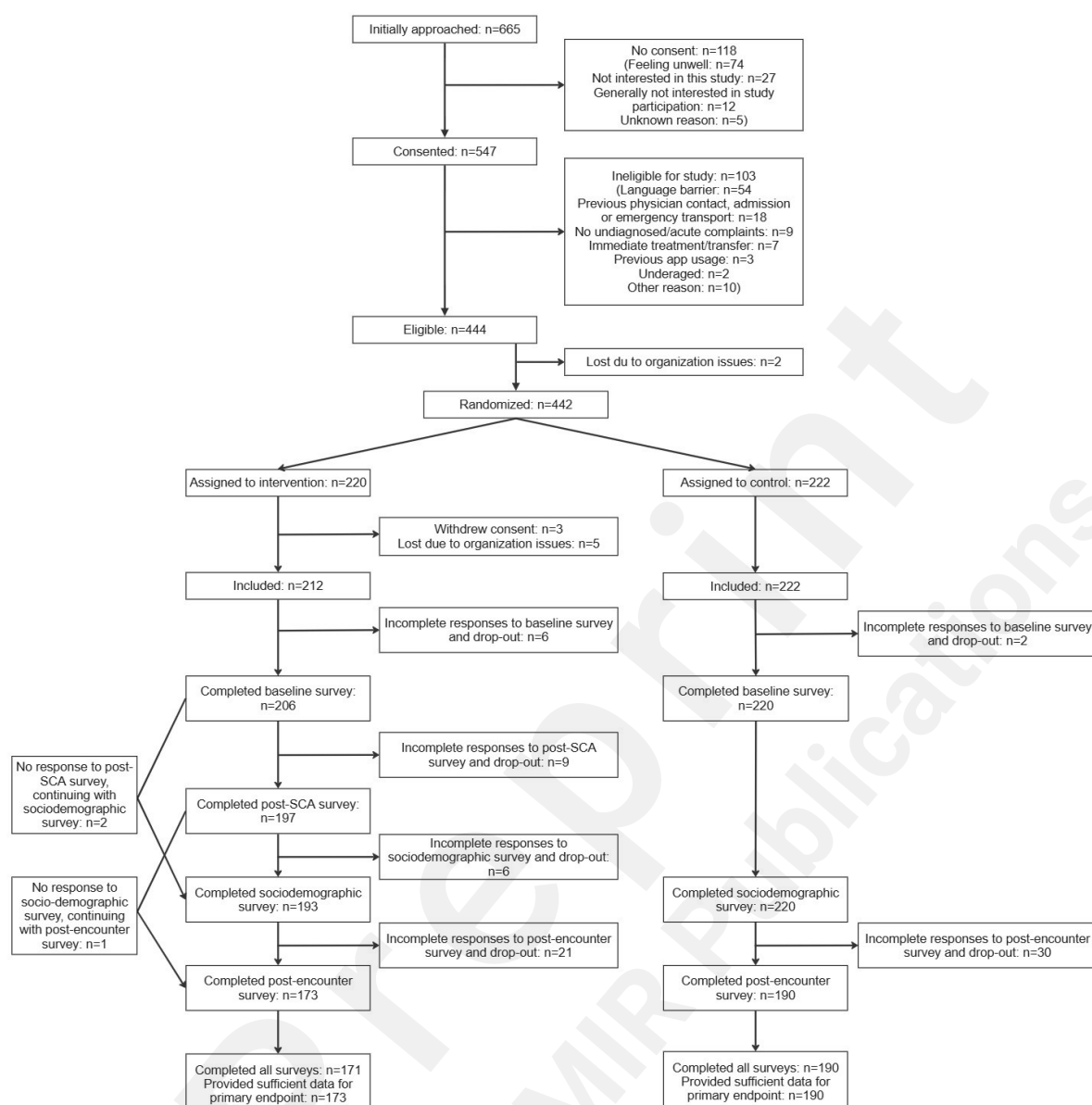
Role of the funding source

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

Results

Recruitment took place from April 11th 2022 to January 25th 2023. We extended the recruitment phase beyond the planned six months to account for staff shortage and COVID-19 pandemic related restrictions. We finished recruitment on the day when we reached the pre-determined sample size. We enrolled 442 participants, 220 of which were randomly assigned to the intervention group, and 222 to the control group. Two eligible participants in the control group dropped out immediately after randomization, one due to organizational issues, the other because of a medical condition. Three participants assigned to the intervention group withdrew their consent and five further participants were lost due to organizational issues. Of the 434 included participants 188 (43.3%) reported a male, 220 (50.7%) female, and 4 (0.9%) diverse sex (the remaining 22 provided no indication of sex). The median age of participants was 36 years (IQR 25). Full details of group characteristics at baseline are reported (Table 1). The baseline survey was completed by 426 participants (206 participants assigned to the intervention and 220 participants to the control group). Of these, 363 provided sufficient data to determine the primary outcome (173 assigned to the intervention group, 190 assigned to the control group), with all of these but two (both assigned to the intervention group) completing all surveys. Figure 1 provides all numbers of participants who completed the surveys over time.

Supplementary Table 1 juxtaposes sex and age distributions of enrolled patients and all patients presenting at these trial sites during the trial period.

Figure 1 Recruitment Flowchart

Trial profile outlining the total number of patients screened, randomized, completing study surveys at multiple times during their stay, and included in the analysis for the primary endpoint. SCA = symptom checker application. We conducted a multicentre, controlled parallel-group trial, at three study sites in Berlin, Germany. Between April 2022 and January 2023, we recruited self-referred, adult walk-in patients presenting with an acute and undiagnosed chief complaint. The intervention group self-assessed their complaints with an SCA prior to the first encounter with a treating physician while the control group received care as usual. After the patient-physician-encounter we surveyed patients and physicians on the patient-physician-interaction and their satisfaction with care.

Table 1 Baseline characteristics

	Control	Intervention
--	---------	--------------

n		222	212
Age in years		36 (19)	30 (16)
Sex			
	Male	96 (43%)	92 (43%)
	Female	122 (55%)	98 (46%)
	Other	2 (1%)	2 (1%)
	NA	2 (1%)	20 (9%)
Questionnaire completed in English		49 (22%)	54 (26%)
Native language			
	German	120 (54%)	103 (49%)
	English	20 (9%)	20 (9%)
	Other	78 (35%)	68 (32%)
	NA	4 (2%)	21 (10%)
Residence			
	Currently residing in Germany	209 (95%)	175 (83%)
	Currently residing outside of Germany	10 (5%)	16 (8%)
	NA	3 (1%)	21 (10%)
Education			
	Student	6 (3%)	7 (3%)
	Basic education	1 (<1%)	0 (0%)
	Lower secondary education after year 9 or 10	13 (6%)	12 (6%)
	Lower secondary education until year 10	38 (17%)	25 (12%)
	High school degree for university entrance	135 (61%)	124 (59%)
	Other school leaving certificate (e.g. awarded abroad)	21 (10%)	23 (11%)
	NA	8 (4%)	21 (10%)
Professional qualification			
	No qualification; still undergoing professional training, e.g. student, trainee, in pre-vocational training, intern	28 (13%)	26 (12%)
	No professional qualification and not undergoing training	12 (5%)	13 (6%)
	Apprenticeship - receiving	21 (9%)	10 (5%)

	vocational training from a company		
	Training at a vocational school, commercial school - combined vocational and academic education	32 (14%)	17 (8%)
	Technical college, e.g. guild school, school for technicians, vocational or professional academy	14 (6%)	11 (5%)
	University of applied sciences, engineering college, advanced technical college	18 (8%)	16 (8%)
	University or college	82 (37%)	87 (41%)
	Other formal qualification (e.g. acquired abroad)	5 (2%)	8 (4%)
	NA	10 (5%)	24 (11%)
Net household income			
	Less than 1.800€	53 (24%)	61 (29%)
	1.801€ to 2.800€	50 (23%)	41 (19%)
	2.801€ to 4.000€	44 (20%)	25 (12%)
	More than 4.000€	39 (18%)	43 (20%)
	NA	36 (16%)	42 (20%)
Anxiety (0-100)		73 (26)	72 (21)
Triage category			
	3	85 (38%)	67 (32%)
	4	104 (47%)	110 (52%)
	5	3 (1%)	1 (<1%)
	NA	30 (14%)	34 (16%)
Self-efficacy (ASKU)			
eHealth literacy (eHEALS)		3.5 (0.88)	3.6 (1)
Self-perceived health			
	Very good	38 (17%)	43 (20%)
	Good	110 (50%)	83 (39%)
	Fair	54 (24%)	47 (22%)
	Bad	14 (6%)	18 (8%)
	Very bad	4 (2%)	1 (<1%)
	NA	2 (1%)	20 (9%)
Chronic morbidity			
	Yes	99 (45%)	74 (35%)
	No	115 (52%)	115 (54%)
	NA	8 (4%)	23 (11%)

Internet usage			
	Daily	201 (91%)	178 (84%)
	Several times a week	15 (7%)	9 (4%)
	Several times a month	4 (2%)	2 (1%)
	Never	0 (0%)	3 (1%)
	NA	2 (1%)	20 (9%)
Tablet usage			
	Daily	37 (17%)	33 (16%)
	Several times a week	32 (14%)	22 (10%)
	Several times a month	24 (11%)	21 (10%)
	Several times a year	36 (16%)	31 (15%)
	Never	86 (39%)	83 (40%)
	NA	7 (3%)	22 (10%)
Previously used SCAs		12 (5%)	14 (7%)

Baseline characteristics of the intention-to-treat population. Data are n, n (%) or median (IQR). ASKU = “Allgemeine Selbstwirksamkeit Kurzskala”, an instrument measuring individuals’ self-efficacy. eHEALS = eHealth Literacy Scale. SCA = symptom checker application. NA refers to both missing responses and respondents indicating their preferences not to provide a response. The higher dropout rate in the intervention group yields a greater proportion of NAs relative to the control group. We conducted a multicentre, controlled parallel-group trial, at three study sites in Berlin, Germany. Between April 2022 and January 2023, we recruited self-referred, adult walk-in patients presenting with an acute and undiagnosed chief complaint. The intervention group self-assessed their complaints with an SCA prior to the first encounter with a treating physician while the control group received care as usual. After the patient-physician-encounter we surveyed patients and physicians on the patient-physician-interaction and their satisfaction with care.

Primary Outcome

Across all three recruiting sites, participants’ mean PSQ scores were close to 80 (on a scale from 0 to 100) in both groups, see Table 2. The linear mixed effects model revealed no significant fixed effect for group (estimate: -2.4, 95%-CI: -6.3 to 1.1, reference: control group). While in two recruiting sites (CCM and JKB) patient-sided PSQ scores were similar between participants in the control and intervention group, for participating patients from the CVK site in the intervention group on average we observed 11-point lower satisfaction with the patient-physician interaction than their controls (see Supplementary Table 2). The employment of different imputation methods revealed no significant effects of missing data (see Supplementary Table 3).

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

	Control	Intervention	p value
Patient satisfaction with patient-physician interaction (patient-sided PSQ)			
Descriptive	80.8 (19.6); n=190	78.5 (20); n=173	..
Estimate for fixed effect of study group in linear mixed model,	..	-2.4 (-6.3 – 1.1)	0.24

intervention to control group			
Patient satisfaction with care (ZUF-8)			
Descriptive	2.6 (0.2); n=190	2.6 (0.2); n=173	..
Estimate for fixed effect of study group in linear mixed model, intervention to control group, 95% CI	..	0.02 (-0.02 – 0.06)	0.27
Change in anxiety level, before SCA use to after			
Descriptive	-	-1.7 (13.5); n=199	..
Estimate for fixed effect of study group in linear mixed model, 95% CI	-	-0.1 (-5.0 – 4.5)	0.96
Participants more anxious after the physician encounter than at baseline			
n (%)	39/191 (20.4%)	36/173 (20.8%)	..
Estimate for fixed effect of study group in generalized linear mixed model, 95% CI	..	0.0 (-0.5 – 0.6)	0.93
Physician satisfaction with patient-physician interaction (physician-sided PSQ)			
Descriptive	76.3 (14.9); n=203	73.7 (15.3); n=191	..
Estimate for fixed effect of study group in linear mixed model, intervention to control group, 95% CI	..	-2.7 (-5.4 – 0.5)	0.08

Summary table of the primary and secondary outcomes. Data are n/N (% or proportion) or mean (SD) unless otherwise stated. PSQ = Patient Satisfaction Questionnaire. ZUF-8 = Fragebogen zur Patientenzufriedenheit, a German version of the Client Satisfaction Questionnaire (CSQ-8). SCA = symptom checker application. Patient anxiety was assessed at baseline, after using the SCA (intervention group only) and after the physician encounter with a visual analogue scale with a range from 0 to 100, with lower values indicating less anxiety. We conducted a multicentre, controlled parallel-group trial, at three study sites in Berlin, Germany. Between April 2022 and January 2023, we recruited self-referred, adult walk-in patients presenting with an acute and undiagnosed chief complaint. The intervention group self-assessed their complaints with an SCA prior to the first encounter with a treating physician while the control group received care as usual. After the patient-physician-encounter we surveyed patients and physicians on the patient-physician-interaction and their satisfaction with care. We surveyed patients' anxiety regarding their symptoms at multiple steps during the trial (before and after the physician encounter, and in the intervention group additionally after using the SCA).

Secondary Outcomes

Satisfaction with care (ZUF-8)

On average, participants in the intervention and control group reported similar ZUF-8 average scores (control

group: mean of average score of 2.6 (SD = 0.2), N = 190; intervention group: mean of average score of 2.6 (SD = 0.2), N = 173). The estimated group difference (0.02, 95%-CI: -0.02 to 0.06, $p = 0.27$) is lower than the hypothesized 0.125.

Anxiety induced by SCA usage

On average, patients reported slightly lower levels of anxiety immediately after using the SCA (66.7 vs. 64.5, i.e. -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, $p = 0.96$). Approximately one third of patients (70/199, 35.2%) in the intervention group reported increased values of anxiety after the SCA assessment, about half ($n=36/70$) of them indicating that their anxiety increased by more than 5 points. At the same time, 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 a fifth of participants reported greater levels of anxiety after the physician encounter than their baseline level (see Table 2).

Treating physicians' satisfaction with interaction (physician-sided PSQ)

Participating physicians completed the physician-facing PSQ in 91% of cases, with scores missing regarding 19/222 patients in the control group and 21/212 patients in the intervention group. Participating physicians on average reported lower PSQ scores (73.7, SD: 15.3) for patients in the intervention group than regarding patients 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) we deemed relevant in the study protocol and was not significant (95%-CI: -5.4 to 0.5, $p = 0.08$). In 2 of the 4 imputation methods we applied, mean differences reached significance (see Supplementary Table 3). These differences still were below the predefined 5-point group-difference we deemed clinically relevant. For all remaining secondary outcomes, none of the four imputation approaches yielded statistically significant differences, see Supplementary Table 3).

Differences in primary and secondary outcome measures between intervention and control group did not turn significant when including only patients of the intervention group whose treating physician took note of the SCA report (data not shown).

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 (57.9%, 95/164). More participants (40.2%, 66/164) reported a (rather) positive influence than a (rather) negative effect (1.8%, 3/164). The numbers were similar when asked about the perceived effect on the care received, see Supplementary Table 4. The numbers also remain similar when considering only cases where the physician indicated to have taken notice of the SCA's report (data not shown).

Effects of the SCA based on physician-reported outcomes

Physicians reported to have taken notice of the SCA's summary report and recommendations for majority of patients in the intervention group (59.9%, 112/187). Physicians most often indicated that the SCA was neither helpful nor unhelpful with the five pre-specified tasks, see Supplementary Table 5). For all five tasks, the SCA was rated (rather) helpful more often than (rather) unhelpful. The SCA was considered most helpful for history taking, diagnosis and conveying information to the patient. This remains unchanged when considering only cases where the treating physician indicated to have seen the SCA's report (data not shown).

Physicians rated their satisfaction with the patient care they delivered, the adequacy of time to diagnosis and

the adequacy of overall length of stay of their patients about equal across both trial groups, see Supplementary Table 6. In both groups, treating physicians reported only very few instances of patients questioning their authority (intervention group: 2/188 (1.1%); control group: 4/184 (2.2%)).

Discussion

The *AkuSym* study was the first RCT to investigate the impact of using an SCA prior to contact with the treating physician on patients' satisfaction with the patient-physician interaction. There was no significant effect of SCA usage in the ED on the primary endpoint (patients' satisfaction with the patient-physician-interaction) or the pre-specified secondary endpoints (patients' anxiety, patients' satisfaction with care, physicians' satisfaction with the patient-physician-interaction). Patients' and physicians' satisfaction with their interaction did not increase with patients' prior use of the SCA, nor did the satisfaction with the care they received or, respectively, the care their physicians provided. Similarly, measures relating to care efficiency, i.e. physicians' appraisal of the time to diagnosis and length of stay, did not demonstrate a benefit of the SCA (no group difference between treatment groups). This contrasts with the perceived effects of the SCA as around a third of patients in the intervention group indicated that the SCA positively impacted their patient-physician-interaction and the care they received.

This discrepancy between measured effects of the SCA and patients' perception thereof is in line with previous literature on patients seeking health information online prior to receiving urgent care services: a survey-based observation study suggested a positive impact with 77.3% (150/196) of patients that had searched their current health problems prior to visiting an ED reported this to have improved the patient-physician relationship.⁵ Meanwhile a RCT primarily investigating the effect of web-based searches on the accuracy of patient generated differential diagnoses did not provide any evidence for an effect of web-based searches on its secondary endpoints (patient-reported satisfaction with care, patients' and physicians' satisfaction with the patient-physician relationship, and patient's anxiety).⁴

We found the same discrepancy between perceived and measurable influence of the SCA on the side of the treating physicians: while the treating physicians of patients in the intervention group often considered their patient's use of the SCA helpful for certain tasks like diagnosis, this did not translate into a relevant difference in the physicians' appraisal of the adequacy of time-to-diagnoses between the intervention and control groups.

These discrepancies give rise to further questions why the perceived positive impact did not materialize in our study: they may, for example, have not been recorded by our choice of endpoints, materialize only in certain subgroups or different usage scenario.

Our trial cannot provide evidence for benefits of SCAs in acute care. However, it also provides no evidence for negative effects of such tools which are commonly at the core of reservations from providers and patients: more patients reported a relevant decrease than increase of anxiety after SCA usage, and only very few physicians reported that their patients questioned their authority - with no difference between the groups.

Two in five physicians did not take notice of the SCA report prior to seeing the respective patient. This might have abated the SCA's effect on the patient-physician-interaction. Leaving it up to the patient to address the report and within the physician's discretion to consult a decision or documentation support tool better reflects the reality of acute care. Thus, we deliberately decided not to recommend physicians to engage with the SCA report.

This study has limitations: SCAs are used in a variety of different scenarios and by users with heterogeneous

expectations.^{1,7,49} Our findings are likely not generalizable to many of these usage scenarios of SCAs. Our patients used the SCA while being in an ED. It is plausible that the SCA might have influenced patients' anxiety levels differently if they had used it at home, outside of a care setting. Our trial included a single SCA of the many available SCAs. Others SCAs may well impact patients, physicians and clinical care differently. We deliberately focused on investigating the SCA's effects on the patient-physician relationship and satisfaction with care. Our study does not address the investigated SCA's utility for navigating patients through a healthcare system safely and efficiently, as we recruited patients who had already decided to seek care in one of the three trial sites. Thus, the participating patients' valuation of the SCA's utility does not relate to the use case of consulting the SCA prior to seeking care.

Further, it is crucial to note that our findings may not reflect important differences between subgroups: The positive and negative effects of SCAs might be specific to a certain subset of users (for example, users with previous experience with online health information, users with hypochondric traits, etc). Most of our patients in the intervention had never used an SCA before. This may be a strength of our trial as it allows us to estimate the effect of SCAs when utilized by a broader and more diverse population than their current users (early adopters). The impact of SCA use might differ between their current users who referred themselves to using these apps, and users prompted to use these apps for example by healthcare providers. However, the proportion of trial participants with prior experience with SCAs is too small in our sample to investigate this meaningfully. Additionally, all our trial sites are in a highly urban location and recruited study participants were younger than the entirety of patients presenting at the trial sites during the recruitment period. Thus, the patient collective is not representative of the entire acute care patient population.

Due to German employee data protection requirements, we were not allowed to match participating patients to their treating physicians. Thus, we are unable to investigate whether physician-related variables confounded our results.

Post-recruitment, we observed a baseline difference in age, with the control group participants being, on average, 5 years older than those in the intervention group. Despite rigorous efforts to identify violations of the study procedure, we found no such violations of the recruitment procedure as the cause. Given that younger age correlated with higher usability reported by the patients in our study, the younger average age of our intervention group could have biased our trial's results in favor of the SCA.

While research is still investigating differences between search-engine based access to online health information and SCAs, advances in generative artificial intelligence (AI) spawned new means for laypersons (and professionals). Though some SCAs, including the one used in this trial, have been shown to have high values for usability, accuracy and safety, our research corroborates that this might not in itself translate into added value for clinical care. This view is encouraged by a recent study not finding a commercially available, patient-facing computerised diagnostic decision support system to improve diagnostic quality⁷³. Even newer generative AI-based applications designed to support clinical care may face a similar struggle⁷⁴. Possible avenues forward are integrating such intelligent support tools in a more sophisticated fashion into existing processes and/or expanding their functions⁷⁵.

Conclusion

In summary, we conducted to our knowledge the first non-industry funded, randomized controlled trial investigating clinically relevant effects of using a SCA (Ada Health) in an acute care setting. Our trial yields no evidence for relevant positive or negative effects of using an SCA prior to the physician encounter on the patient-physician relationship and satisfaction with care. Both patients and physicians more often perceived

the SCA's influence as positive than negative. Thus, it remains an open question whether the perceived positive effects are unsubstantiated, were not captured by our choice of endpoints, or materialize only in certain subgroups or different usage scenarios. Our study did not reveal negative effects of SCA use commonly described in the literature such as inducing anxiety or eroding patients' trust in healthcare professionals. Our trial underscores the need for and value of clinical trials on mobile health applications as high usability and high levels of reported accuracy did not translate in improving patient and physician reported outcomes.

Contributors

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 analysed 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.

Data Sharing

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

Declaration of Interests

HN: Invests in exchange-traded funds that hold shares of various companies in the technology and healthcare sectors that do not gain financially from the publication of the manuscript but might do so in some way in the future. He is a member of the Digital Health Working group of the German Network Health Service Research (DNVF), the section "Digitalization" of the German College of General Practitioners and Family Physicians (DEGAM). Hendrik Napierala reports grants from the Deutsche Forschungsgemeinschaft (German Research Council, DFG), the Innovationsfonds des Gemeinsamen Bundesausschuss (Federal Joint Committee of Germany), the German Ministry of Research and Education (BMBF) and the German Ministry of Health (BMG).

MB: Reports grants from Bundesministerium für Gesundheit (German Ministry of Health).

MK: Invests in exchange-traded funds that hold shares of various companies in the technology and healthcare sectors. He is also holding stocks in organizations that do not gain financially from the publication of the manuscript but might do so in some way in the future. Reports grants from the Berlin University Alliance.

MBA: Invests in exchange-traded funds that hold shares of various companies in the technology and healthcare sectors that do not gain financially from the publication of the manuscript but might do so in some way in the future.

MB: Myrto Bolanaki reports grants from Bundesministerium für Gesundheit (German Ministry of Health) during the conduct of the study, grants from Bundesministerium für Bildung und Forschung (German Ministry for Education and Research), grants from Thermo Fisher BRAHMS, grants from Roche Diagnostics outside the submitted work.

KS: Konrad Schmidt reports grants from the Deutsche Forschungsgemeinschaft (German Research Council) and the Innovationsfonds des Gemeinsamen Bundesausschuss (Federal Joint Committee of Germany).

LS: Invests in exchange-traded funds that hold shares of various companies in the technology and healthcare sectors that do not gain financially from the publication of the manuscript but might do so in some way in the future.

CW: reports grants from the Deutsche Forschungsgemeinschaft (German Research Council) and the Innovationsfonds des Gemeinsamen Bundesausschuss (Federal Joint Committee of Germany).

SKP: None.

CH: Christoph Heintze reports grants from the Bundesministerium für Bildung und Forschung (German Ministry for Education and Research), Deutsche Forschungsgemeinschaft (German Research Council) and the Innovationsfonds des Gemeinsamen Bundesausschuss (Federal Joint Committee of Germany).

MM: M. Möckel has received grants from Health Care Research Projects and Biomarker Research and personal fees from Consulting outside the submitted work.

FB: F. Balzer reports grants from the German Federal Ministry of Education and Research, grants from the German Federal Ministry of Health, grants from Berlin Institute of Health, personal fees from Elsevier Publishing, grants from Hans Böckler Foundation, others from Robert Koch Institute, grants from Einstein Foundation, grants from Berlin University Alliance, personal fees from Medtronic, outside the submitted work.

AS: Prof. Slagman reports grants from Bundesministerium für Gesundheit (German Ministry of Health), during the conduct of the study. Grants from Deutsche Forschungsgemeinschaft (German Research Council), grants from Bundesministerium für Bildung und Forschung (German Ministry for Education and Research), grants from Innovationsfonds des Gemeinsamen Bundesausschuss (Federal Joint Committee of Germany), grants from Zentralinstitut für die Kassenärztliche Versorgung (Research Institute of Statutory Health Insurance Physicians in Germany), grants from Thermo Fisher Scientific, grants from Roche Diagnostics, personal fees from Biomarkers - Scientific Journal (Associate Editor), outside the submitted work.

MLS: Invests in exchange-traded funds that hold shares of various companies in the technology and healthcare sectors that do not gain financially from the publication of the manuscript but might do so in some way in the future. In 2014 and 2015 he was an employee at Ada Health GmbH (formerly known as medx GmbH), the developer company of the investigated health application.

Acknowledgments

This study was supported by grants from the German Ministry of Health (Funding number 2521TEL500).

Supplementary Material

Supplementary Appendix

References

- 1 One in two EU citizens look for health information online. <https://ec.europa.eu/eurostat/web/products-eurostat-news/-/edn-20210406-1> (accessed March 12, 2024).
- 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**: 617–25.
- 3 Baumann E, Czerwinski F, Rosset M, Seelig M, Suhr R. Wie informieren sich die Menschen in Deutschland zum Thema Gesundheit? Erkenntnisse aus der ersten Welle von HINTS Germany. *Bundesgesundheitsbl* 2020; **63**: 1151–60.
- 4 Martin SS, Quaye E, Schultz S, *et al.* A randomized controlled trial of online symptom searching to inform patient generated differential diagnoses. *npj Digit Med* 2019; **2**: 110.
- 5 Cocco AM, Zordan R, Taylor DM, *et al.* Dr Google in the ED: searching for online health information by adult emergency department patients. *Medical Journal of Australia* 2018; **209**: 342–7.
- 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* 2021; **7**: e22637.
- 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* 2020; **22**: e14679.
- 8 Kao C-K, Liebovitz DM. Consumer Mobile Health Apps: Current State, Barriers, and Future Directions. *PM R* 2017; **9**: S106–15.
- 9 Graafen MA, Sennekamp M, Messemaker A. „Also, im Internet steht ...“ – Wenn Ärztinnen und Ärzte auf internetinformierte Patienten treffen. *ZFA* 2021; **97**: 210–4.
- 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* 2020; **20**: 848.
- 11 Preiser C, Radionova N, Ög E, *et al.* The Doctors, Their Patients, and the Symptom Checker App: Qualitative Interview Study With General Practitioners in Germany. *JMIR Hum Factors* 2024; **11**: e57360.
- 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* 2020; **270**: 966–70.
- 13 Doherty-Torstrick ER, Walton KE, Fallon BA. Cyberchondria: Parsing Health Anxiety From Online Behavior. *Psychosomatics* 2016; **57**: 390–400.
- 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. *Br J Gen Pract Open* 2017; **1**: BJGP-2017-0833.
- 15 Using technology to ease the burden on primary care. Healthwatch Enfield, 2019 https://www.healthwatch.co.uk/sites/healthwatch.co.uk/files/reports-library/20190122_Enfield_%20Using%20technology%20to%20ease%20the%20burden%20on%20primary%20care.pdf (accessed March 13, 2021).
- 16 Wetzel A-J, 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). *DIGITAL HEALTH* 2024; **10**: 20552076241231555.
- 17 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**: 17.
- 18 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.
- 19 EPatient Analytics GmbH. EPatient Survey 2020. 2020; published online Nov 3. <https://www.hcm-magazin.de/epatient-survey-2020-digital-health-studie/150/10992/407743> (accessed March 6, 2021).
- 20 Kassenärztliche Bundesvereinigung. Patienten-Navi. 2025; published online Jan 18. <https://www.116117.de/de/patienten-navi.php>.
- 21 NHS. NHS 111 Online. 2025; published online Jan 18. <https://111.nhs.uk/>.
- 22 Sutter Health. Sutter Health Teams Up With Ada Health to Improve Patient Care by Delivering On-Demand Healthcare Guidance. 2019; published online Feb 13. <https://news.sutterhealth.org/2019/02/11/7021/> (accessed Nov 13, 2021).
- 23 Kaiser Permanente Health Plan, Inc. Check Your Symptoms. <https://healthy.kaiserpermanente.org/health-wellness/health-encyclopedia/he.check-your-symptoms.hwsxchk> (accessed Nov 13, 2021).
- 24 Sana Kliniken AG. Symptom Checker in Kooperation mit Infermedica. 2025; published online Jan 18. <https://www.sana.de/unternehmen/digital/digitale-loesungen-im-einsatz/symptom-checker>.
- 25 Verband der Privaten Krankenversicherung e.V. KI-Spezialist Infermedica expandiert in Deutschland. 2024; published online Feb 14. <https://www.pkv.de/verband/presse/meldungen/ki-spezialist-infermedica-gelingt-eintritt-in-den-gkv-markt/>.
- 26 Hill MG, Sim M, Mills B. The quality of diagnosis and triage advice provided by free online symptom checkers and apps in Australia. *Medical Journal of Australia* 2020; : mja2.50600.
- 27 Semigran HL, Linder JA, Gidengil C, Mehrotra A. Evaluation of symptom checkers for self

- diagnosis and triage: audit study. *BMJ* 2015; : h3480.
- 28Ceney 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**: e0254088.
- 29Hennemann 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* 2022; **9**: e32832.
- 30Knitz J, Hasanaj R, Beyer J, 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* 2024; **26**: e56514.
- 31Knitz J, Tascilar K, Fuchs F, 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* 2024; **26**: e55542.
- 32Fraser 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* 2023; **11**: e49995.
- 33Schmieding 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**: e31810.
- 34Riboli-Sasco E, El-Osta A, Alaa A, et al. Triage and Diagnostic Accuracy of Online Symptom Checkers: Systematic Review. *J Med Internet Res* 2023; **25**: e43803.
- 35Meer A, Rahm P, Schwendinger M, 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* 2024; **26**: e58157.
- 36Hammoud M, Douglas S, Darmach M, Alawneh S, Sanyal S, Kanbour Y. Evaluating the Diagnostic Performance of Symptom Checkers: Clinical Vignette Study. *JMIR AI* 2024; **3**: e46875.
- 37Liu 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* 2024; **11**: e55099.
- 38Verzantvoort 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**: e0199284.
- 39Winn 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**: e1918561.
- 40Chambers D, Cantrell A, Johnson M, et al. Digital and online symptom checkers and assessment services for urgent care to inform a new digital platform: a systematic review. *Health Serv Deliv Res* 2019; **7**: 1–88.
- 41Gellert GA, Orzechowski PM, Price T, et al. A multinational survey of patient utilization of and

- value conveyed through virtual symptom triage and healthcare referral. *Front Public Health* 2023; **10**: 1047291.
- 42 Gellert GA, Kabat-Karabon A, Gellert GL, *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.
- 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* 2019; **14**: 69.
- 44 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**: e38364.
- 45 Gottlieb K, Petersson G. Limited evidence of benefits of patient operated intelligent primary care triage tools: findings of a literature review. *BMJ Health Care Inform* 2020; **27**: e100114.
- 46 Pairen 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.
- 47 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**: 118.
- 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. *Journal of Medical Internet Research* 2023; **25**: e39219.
- 49 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**: 17.
- 50 Snow G. Blockrand: Randomization for block random clinical trials. 2020. <https://cran.r-project.org/web/packages/blockrand/index.html>.
- 51 Wessel D, Attig C, Franke T. ATI-S - An Ultra-Short Scale for Assessing Affinity for Technology Interaction in User Studies. In: Proceedings of Mensch und Computer 2019. Hamburg Germany: ACM, 2019: 147–54.
- 52 Robine J-M, Jagger C. Report to Eurostat on European Health Status Module. *Montpellier: Euro-REVES* 2003.
- 53 Beierlein C, Kovaleva A, Kemper CJ, Rammstedt B. Allgemeine Selbstwirksamkeit Kurzskala (ASKU). *Zusammenstellung sozialwissenschaftlicher Items und Skalen (ZIS)* 2014. DOI:10.6102/ZIS35.
- 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.
- 55 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.

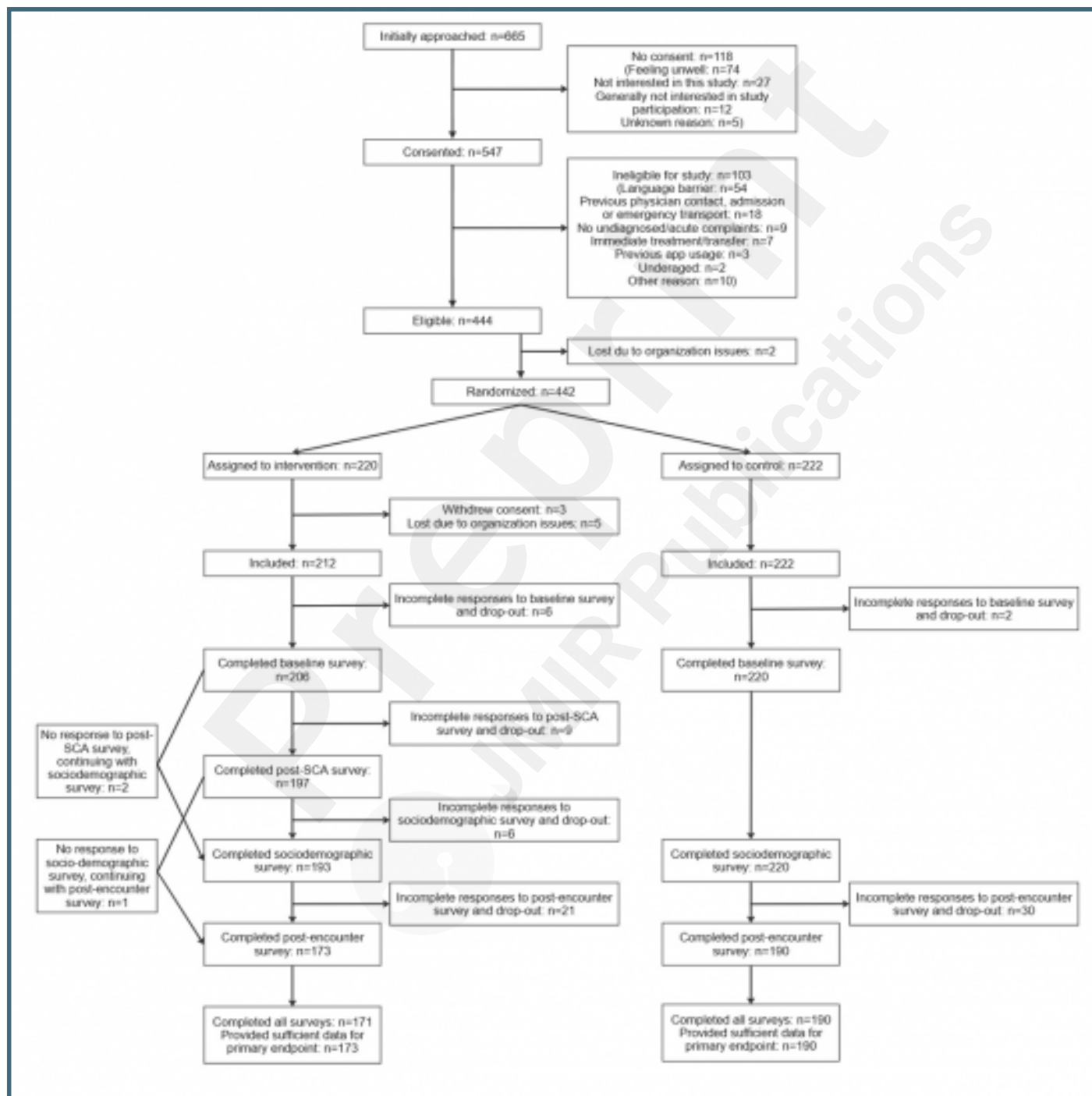
- 56Mehl A, Bergey F, Cawley C, Gilsdorf A. Syndromic Surveillance Insights from a Symptom Assessment App Before and During COVID-19 Measures in Germany and the United Kingdom: Results From Repeated Cross-Sectional Analyses. *JMIR Mhealth Uhealth* 2020; **8**: e21364.
- 57Ćirković A. Evaluation of Four Artificial Intelligence–Assisted Self-Diagnosis Apps on Three Diagnoses: Two-Year Follow-Up Study. *J Med Internet Res* 2020; **22**: e18097.
- 58Cotte F, Mueller T, Gilbert S, *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* 2022; **10**: e32340.
- 59Morse KE, Ostberg NP, Jones VG, Chan AS. Digital Symptom Checker Usage and Triage: Population-Based Descriptive Study in a Large North American Integrated Health System. 2020; published online Nov 7. DOI:10.2196/20549.
- 60Gilbert 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 Primary Health* 2021; **27**: 377–81.
- 61Blanchard CG, Ruckdeschel JC, Fletcher BA, Blanchard EB. The impact of oncologists' behaviors on patient satisfaction with morning rounds. *Cancer* 1986; **58**: 387–93.
- 62Zandbelt 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* 2004; **19**: 1088–95.
- 63Schmidt J, Lamprecht F, Wittmann WW. Zufriedenheit mit der stationären Versorgung. Entwicklung eines Fragebogens und erste Validitätsuntersuchungen. [Satisfaction with inpatient care: Development of a questionnaire and first validity assessments.]. *PPmP: Psychotherapie Psychosomatik Medizinische Psychologie* 1989; **39**: 248–55.
- 64Attkisson 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, US: Lawrence Erlbaum Associates, Inc, 1994: 402–20.
- 65Abend R, Dan O, Maoz K, Raz S, Bar-Haim Y. Reliability, validity and sensitivity of a computerized visual analog scale measuring state anxiety. *Journal of Behavior Therapy and Experimental Psychiatry* 2014; **45**: 447–53.
- 66R Core Team. R: A Language and Environment for Statistical Computing. 2024; published online April 24. <https://www.R-project.org/>.
- 67Wickham H, Averick M, Bryan J, *et al.* Welcome to the Tidyverse. *JOSS* 2019; **4**: 1686.
- 68Bates D, Mächler M, Bolker B, Walker S. Fitting Linear Mixed-Effects Models Using **lme4**. *J Stat Soft* 2015; **67**. DOI:10.18637/jss.v067.i01.
- 69Lüdecke D, Ben-Shachar M, Patil I, Makowski D. Extracting, Computing and Exploring the Parameters of Statistical Models using R. *JOSS* 2020; **5**: 2445.
- 70Buuren SV, Groothuis-Oudshoorn K. **mice**: Multivariate Imputation by Chained Equations in R. *J Stat Soft* 2011; **45**. DOI:10.18637/jss.v045.i03.

- 71 Napierala H, Kopka M, Altendorf MB, *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* 2022; **23**: 791.
- 72 Schulz KF, Altman DG, Moher D, CONSORT Group. CONSORT 2010 statement: updated guidelines for reporting parallel group randomised trials. *BMJ* 2010; **340**: c332.
- 73 Hautz WE, Marcin T, Hautz SC, *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. *The Lancet Digital Health* 2025; **7**: e136–44.
- 74 Goh E, Gallo R, Hom J, *et al.* Large Language Model Influence on Diagnostic Reasoning: A Randomized Clinical Trial. *JAMA Netw Open* 2024; **7**: e2440969.
- 75 Kostopoulou O, Delaney B. AI for medical diagnosis: does a single negative trial mean it is ineffective? *The Lancet Digital Health* 2025; **7**: e108–9.

Supplementary Files

Figures

Trial profile outlining the total number of patients screened, randomized, completing study surveys at multiple times during their stay, and included in the analysis for the primary endpoint. SCA = symptom checker application. We conducted a multicentre, controlled parallel-group trial, at three study sites in Berlin, Germany. Between April 2022 and January 2023, we recruited self-referred, adult walk-in patients presenting with an acute and undiagnosed chief complaint. The intervention group self-assessed their complaints with an SCA prior to the first encounter with a treating physician while the control group received care as usual. After the patient-physician-encounter we surveyed patients and physicians on the patient-physician-interaction and their satisfaction with care.



Multimedia Appendixes

Supplementary Tables.

URL: <http://asset.jmir.pub/assets/8948dbe5ecea3ded5a838f976ff0d7ff.docx>



CONSORT (or other) checklists

Consort Checklist for the AkuSym RCT.

URL: <http://asset.jmir.pub/assets/4e326e485e702c8c544b838b9b975231.pdf>