

Original Paper

A Symptom-Checker for Adult Patients Visiting an Interdisciplinary Emergency Care Center and the Safety of Patient Self-Triage: Real-Life Prospective Evaluation

Andreas Meer¹, MSc, MD; Philipp Rahm², MD; Markus Schwendinger², MD; Michael Vock³, PhD; Bettina Grunder¹, MD; Jacopo Demurtas¹, MD, PhD; Jonas Rutishauser⁴, MD

¹In4medicine Inc, Bern, Switzerland

²Cantonal Hospital Baden, Baden, Switzerland

³Institute of Mathematical Statistics and Actuarial Science, University of Bern, Bern, Switzerland

⁴Clinical Trial Unit, Cantonal Hospital Baden and Medical Faculty, University of Basel, Baden, Switzerland

Corresponding Author:

Andreas Meer, MSc, MD

In4medicine Inc

Monbijoustrasse 23

Bern, 3011

Switzerland

Phone: 41 313701330

Email: a.meer@in4medicine.ch

Abstract

Background: Symptom-checkers have become important tools for self-triage, assisting patients to determine the urgency of medical care. To be safe and effective, these tools must be validated, particularly to avoid potentially hazardous undertriage without leading to inefficient overtriage. Only limited safety data from studies including small sample sizes have been available so far.

Objective: The objective of our study was to prospectively investigate the safety of patients' self-triage in a large patient sample. We used SMASS (Swiss Medical Assessment System; in4medicine, Inc) pathfinder, a symptom-checker based on a computerized transparent neural network.

Methods: We recruited 2543 patients into this single-center, prospective clinical trial conducted at the cantonal hospital of Baden, Switzerland. Patients with an Emergency Severity Index of 1-2 were treated by the team of the emergency department, while those with an index of 3-5 were seen at the walk-in clinic by general physicians. We compared the triage recommendation obtained by the patients' self-triage with the assessment of clinical urgency made by 3 successive interdisciplinary panels of physicians (panels A, B, and C). Using the Clopper-Pearson CI, we assumed that to confirm the symptom-checkers' safety, the upper confidence bound for the probability of a potentially hazardous undertriage should lie below 1%. A potentially hazardous undertriage was defined as a triage in which either all (consensus criterion) or the majority (majority criterion) of the experts of the last panel (panel C) rated the triage of the symptom-checker to be "rather likely" or "likely" life-threatening or harmful.

Results: Of the 2543 patients, 1227 (48.25%) were female and 1316 (51.75%) male. None of the patients reached the prespecified consensus criterion for a potentially hazardous undertriage. This resulted in an upper 95% confidence bound of 0.1184%. Further, 4 cases met the majority criterion. This resulted in an upper 95% confidence bound for the probability of a potentially hazardous undertriage of 0.3616%. The 2-sided 95% Clopper-Pearson CI for the probability of overtriage (n=450 cases, 17.69%) was 16.23% to 19.24%, which is considerably lower than the figures reported in the literature.

Conclusions: The symptom-checker proved to be a safe triage tool, avoiding potentially hazardous undertriage in a real-life clinical setting of emergency consultations at a walk-in clinic or emergency department without causing undesirable overtriage. Our data suggest the symptom-checker may be safely used in clinical routine.

Trial Registration: ClinicalTrials.gov NCT04055298; <https://clinicaltrials.gov/study/NCT04055298>

(*J Med Internet Res* 2024;26:e58157) doi: [10.2196/58157](https://doi.org/10.2196/58157)

KEYWORDS

safety; telemedicine; teletriage; symptom-checker; self-triage; self-assessment; triage; triaging; symptom; symptoms; validation; validity; telehealth; mHealth; mobile health; app; apps; application; applications; diagnosis; diagnoses; diagnostic; diagnostics; checker; checkers; check; web; neural network; neural networks

Introduction

In potentially critical situations, clinical warning signs and symptoms may be considered too late by patients due to a lack of professional triage [1]. In this context, various initiatives have been launched to improve outpatient emergency care and the population's access to a low-threshold initial medical assessment [2]. Symptom-checkers, which enable medical self-triage, have recently been introduced for this purpose. Such tools could assist the increasing number of persons without ready access to a primary care physician, for example, migrants or young persons who had been previously healthy. If implemented in settings outside the hospital, that is, at home or work, tools for efficient and safe self-triage could help avoid unnecessary emergency hospital visits, thus contributing to reducing overcrowding and costs.

To fulfill the regulatory requirements [3] and to be used as a part of standard care, the appropriateness and safety of these instruments must be evaluated in concrete clinical settings with real patients [4-6].

Appropriate care results from adequate triage and treatment, while inappropriate care may lead to unsuitable or even dangerous health care delivery. The concept of appropriateness hence includes a widespread range of quality aspects, of which safety is only one. The difficulty of assessing appropriateness in health care and of gaining agreement between clinicians on acceptable and safe care is highlighted by different authors [7,8]. When assessing the appropriateness of medical triage, the question "Was the decision right" suggests that there is merely one single correct triage decision. This question does not appropriately reflect the complex interaction of clinical, social, and environmental factors in medical decision-making. Rather, physicians should consider a range of appropriate triage decisions to guide their actions. Safety is an essential quality attribute of a medical service. In contrast to the idea of appropriateness, the concept of safety focuses on the risk of a specific conduct. When asking about the safety of a symptom-checker, a risk-based approach should be taken, and safety should encompass possible risks to a patient's health and life [3].

An evaluation of 23 symptom-checkers using 45 patient vignettes concluded that most symptom-checkers were deficient in both appropriate triage and correct diagnosis [9]. However, the study did not comment on the safety of the tested devices. A review paper including 14 studies found inconsistent evidence regarding the triage and diagnostic appropriateness of symptom-checkers for common health problems. The average appropriateness of triage ranged from 27% to 92%. This paper did not specifically evaluate the safety of symptom-checkers [5]. Another review paper cited only 6 studies that analyzed the safety of symptom-checkers [10]. These studies were mostly

short-term and included samples that were too small and heterogeneous to make reliable statements about safety.

Given the present shortage of data on self-triage, we aimed to investigate the safety of a newly developed symptom-checker (SMASS; Swiss Medical Assessment System) in a concrete clinical setting with patients seeking emergency care.

Methods

Study Design

Before the inclusion of the first patient, this study was registered (ClinicalTrials.gov identifier: NCT04055298).

This study was performed between November 25, 2019, and May 1, 2020, at the walk-in clinic and interdisciplinary emergency department (WIC/ED) of the cantonal hospital of Baden, Switzerland. The WIC/ED is open 24 hours a day, 365 days a year and treats about 55,000 patients annually. Patients are routinely triaged by a nurse using the Emergency Severity Index (ESI) [11]. ESI 1-2 patients are treated in the ED, while ESI 3-5 patients are treated in the WIC.

The symptom-checker used in this study (SMASS in the pathfinder version, release 4.1.12) was developed by in4medicine, Inc. The first author (author AM) is the chief executive officer and founder of this company. To minimize bias, a majority of independent researchers were involved in this study, including establishing the protocol and all practical aspects of the trial. No employee of in4medicine took part in the actual conduct of the trial. Data analysis and statistical calculations were performed by an independent biostatistician. This study was independently monitored by the clinical trial unit of the Medical Faculty of the University of Bern.

The SMASS pathfinder symptom-checker is a medical device class I under the Medical Device Directive and medical device class IIb under the Medical Device Regulation. The Conformité Européenne declaration of conformity to the Swiss Agency for Therapeutic Products (Swissmedic) was made on June 4, 2018. The symptom-checker is a web-based software that aims to support health professionals and laypersons in the structured documentation and assessment of health problems and to advise users about possible medical assessment steps and treatment measures. It is based on a computerized neural network that incorporates extensive data from scientific studies, guidelines, and expertise from various professional boards of specialists in the field of prehospital medical triage. The symptom-checker provides digitalized questionnaires of 125 frequent reasons for consultations (eg, fever, cough, and abdominal pain) and their associated red flags. Based on the triage result, a report including patient gender, age group, symptoms, medical history, and recommendations as to the appropriate time-to-treat and point-of-care is provided. Depending on the presence of red flags, the symptom-checker assigns the clinical condition of the patient to a triage level (Tables 1 and 2). If 5 or more assessment

questions are answered as “unclear,” the user is notified that the software cannot provide targeted triage advice and that the patient should seek immediate consultation with a physician concerning his or her medical complaints.

Table 1. Triage levels as recommended to the patient by the symptom-checker. Recommendations are given regarding time-to-treat (emergency, immediately, today, later, and unclear) and point-of-care (ambulance, hospital, physician, call center, pharmacy, self-care, and unclear).

	Ambulance	Hospital	Doctor	Call center	Pharmacy	Self-care	Unclear
Emergency	16	15	— ^a	—	—	—	—
Immediately	14	13	12	—	—	—	—
Today	—	11	10	8	6	4	—
Later	—	—	9	7	5	3	1
Unclear	—	—	—	—	—	2	0

^aNot applicable.

Table 2. Recommended actions to be taken by the patient, as defined by triage levels (left column). Levels range from 0 (lowest level) to 16 (highest level). Interpretations of the triage level and measures to be taken are specified in the right column.

Triage level	Name	Recommended action
Level 16	Emergency ambulance	CPR ^a /CPR readiness. There is a potentially life-threatening condition. Medical treatment must be given now. Alert the emergency services via the number 144.
Level 15	Emergency hospital	CPR/CPR readiness. There is a potentially life-threatening condition. Medical treatment must be given now. Alert the emergency services via the number 144. Medical treatment should be provided at a hospital.
Level 14	Immediately ambulance	Medical treatment does not allow any delay. Treatment should be given immediately. Alert the emergency services via the number 144.
Level 13	Immediately hospital	Medical treatment does not allow any delay. Treatment should be given immediately. Medical treatment should be provided at a hospital.
Level 12	Immediately doctor	Medical treatment does not allow any delay. Treatment should be given immediately. Medical treatment should be provided by a registered doctor ^b .
Level 11	Today hospital	Medical treatment does not have to take place immediately, but should not be delayed until tomorrow or over the weekend. Medical treatment should take place within the next 24 hours. Medical treatment should be provided at a hospital.
Level 10	Today doctor	Medical treatment does not have to take place immediately, but should not be delayed until tomorrow or over the weekend. Medical treatment should take place within the next 24 hours. Medical treatment should be provided by a registered doctor ^b .
Level 9	Later doctor	Medical treatment is not urgent. If the symptoms do not subside in the next 2 days, treatment by a doctor is indicated. Medical treatment should be provided by a registered doctor ^b .
Level 8	Today call center	Medical treatment does not have to take place immediately, but should not be delayed until tomorrow or over the weekend. Medical treatment should take place within the next 24 hours. The affected person should be advised by a telemedicine center on how to proceed.
Level 7	Later call center	Medical treatment is not urgent. If the symptoms do not subside in the next 2 days, treatment by a doctor is indicated. The affected person should be advised by a telemedicine center on how to proceed.
Level 6	Today pharmacy	Medical treatment does not have to take place immediately, but should not be delayed until tomorrow or over the weekend. Medical treatment should take place within the next 24 hours. The affected person should be advised at a pharmacy on how to proceed.
Level 5	Later pharmacy	Medical treatment is not urgent. If the symptoms do not subside in the next 2 days, treatment by a doctor is indicated. The affected person should be advised at a pharmacy on how to proceed.
Level 4	Today self-care	Medical treatment does not have to take place immediately, but should not be delayed until tomorrow or over the weekend. Medical treatment should take place within the next 24 hours. The complaints can be treated independently by simple measures.
Level 3	Later self-care	Medical treatment is not urgent. If the symptoms do not subside in the next 2 days, treatment by a doctor is indicated. The complaints can be treated independently by simple measures.
Level 0-2	Unclear	The survey contains too many ambiguities. A targeted initial assessment is not possible.

^aCPR: cardiopulmonary resuscitation.

^bFor example, family doctor, family doctor substitute, family doctor emergency service, or suitable specialist.

All patients aged ≥ 18 years attending the WIC/ED between 8 AM and 5 PM were eligible. Exclusion criteria included aged < 18 years; ESI 1 patients requiring immediate, life-saving intervention; inability to use a tablet PC; inability to communicate in German, French, Italian, or English; inability or unwillingness to give written informed consent and follow the procedures of this study; known or suspected noncompliance; known drug or alcohol abuse; the presence of symptoms or complaints not encompassed by the symptom-checker database (eg, long-lasting hiccups, hair loss).

After instruction by the study's staff and providing written informed consent, the participants independently assessed their health status and complaints as instructed by the

symptom-checker on a tablet PC. They were subsequently evaluated and treated by routine medical staff.

In primary care, medical triage decisions usually have to be based solely on the patient's symptoms. We have chosen experts independent of the treatment (panels A, B, and C) as evaluators to ensure that the triage decision is based purely on the symptoms of this study's patients. Including treating physicians as comparators in this study could have influenced the triage decision by additional information (physical examination and diagnostic test results).

Our evaluation of the symptom-checker focused on safety, as this is an essential quality attribute of a medical device [3]. To reflect the highly individual nature of medical decision-making,

which usually results in low interrater reliability [12-14], an independent team of experienced physicians engaged in a stepwise evaluation procedure in which each case that was classified as undertriaged by panel A experts was assessed by several experts.

A research assistant and 3 external interdisciplinary panels of board-certified physicians were involved in the evaluation process (panel A, 5 experts; panel B, 2 experts; panel C, 5 experts). Except for one of the 12 panelists (author BG), they were not affiliated with in4medicine, Inc. None of them took part in the conduct of this study. For every patient, the symptom-checker issued a report summarizing the clinical information. Patients and panelists were unaware of the triage-level recommendations (time-to-treat and point-of-care) made by the symptom-checker. All reports were first assessed by members of panel A, who adjudicated an appropriate range of triage levels to every case. The research assistant then compared the adjudication of the panel A experts with the recommendation issued by the symptom-checker. If the comparison showed that the recommendation of the symptom-checker was below the appropriate range of triage levels determined by the rater of panel A, hence was undertriaged, the case was assigned to panel B. In 80 instances, panelists erroneously examined the same cases twice and concluded on diverse triage recommendations. In these cases, the first of the 2 recommendations was used for the analysis.

The evaluation procedure was repeated by panel B. Each of the 2 panelists evaluated all diverging cases. If the case was undertriaged according to 3 experts (1 expert from panel A and 2 experts from panel B), the case was subsequently analyzed by panel C.

Each member of panel C individually assessed the clinical safety of the triage decision based on the complete structured reports generated by the symptom-checker as well as the WIC/ED's redacted discharge reports. Each of the 5 panelists decided individually on potentially hazardous undertriage. In a modified Delphi process, the panelists first individually adjudicated potentially hazardous undertriage on a 4-point Likert scale. Possible ratings were "unlikely," "rather unlikely," "rather likely," and "likely" that the patient was exposed to a risk to life or health. If the panelists subsequently reached a consensus that the triage of the symptom-checker was "rather likely" or "likely" exposing a patient to a risk to life or health, the case was considered a potentially hazardous undertriage (consensus criterion). As a complement to the original analysis plan, a modified criterion for potentially hazardous undertriage was

evaluated, defined as a majority of panel C members judging a risk to life or health as "rather likely" or "likely" (majority criterion).

The primary analysis consisted of the calculation of the 95% upper Clopper-Pearson confidence bound for the probability of undertriage resulting in a risk to life or health (potentially hazardous undertriage). To confirm the safety of the symptom-checker, this upper confidence bound should lie below 1%. For the sample size calculation, we assumed that a 20% probability of failure to meet this criterion is acceptable for a true probability of potentially hazardous undertriage of no more than 0.5%. This is equivalent to requiring a 1-sided test at level 5% to show that the probability of potentially hazardous undertriage is below 1% with a power of 80%, assuming that the true probability is 0.5%. This resulted in a minimal sample size of 2185 patients. Accounting for an estimated rate of 2% "unclear" responses, at least 2230 patients were planned to be included. Secondary analyses included central 95% Clopper-Pearson CIs for the further probabilities, based on corresponding empirical proportions. The software R (version 4.2.0; R Foundation for Statistical Computing) was used for the statistical evaluations.

Ethical Considerations

This study was approved by the competent ethics committee (Ethikkommission Nordwest- und Zentralschweiz EKNZ, project ID 01784) and was conducted per the most recent version of the Declaration of Helsinki, complying with International Council for Harmonisation of Technical Requirements for Pharmaceuticals for Human Use—good clinical practice and International Organization for Standardization European Norm 14155 (clinical investigation of medical devices for human subjects—good clinical practice) as well as with applying national legal and regulatory requirements. All patients gave written informed consent to participate in this study. They did not receive any financial or other compensation. Patients were anonymized upon data collection. Discharge notes studied by panel C were redacted.

Generative artificial intelligence was not used in any portion of this paper's writing.

Results

The baseline characteristics of the participants are shown in [Table 3](#) and the recommendations obtained by the symptom-checker in [Table 4](#). [Figure 1](#) shows the flow of analyses by panels A-C.

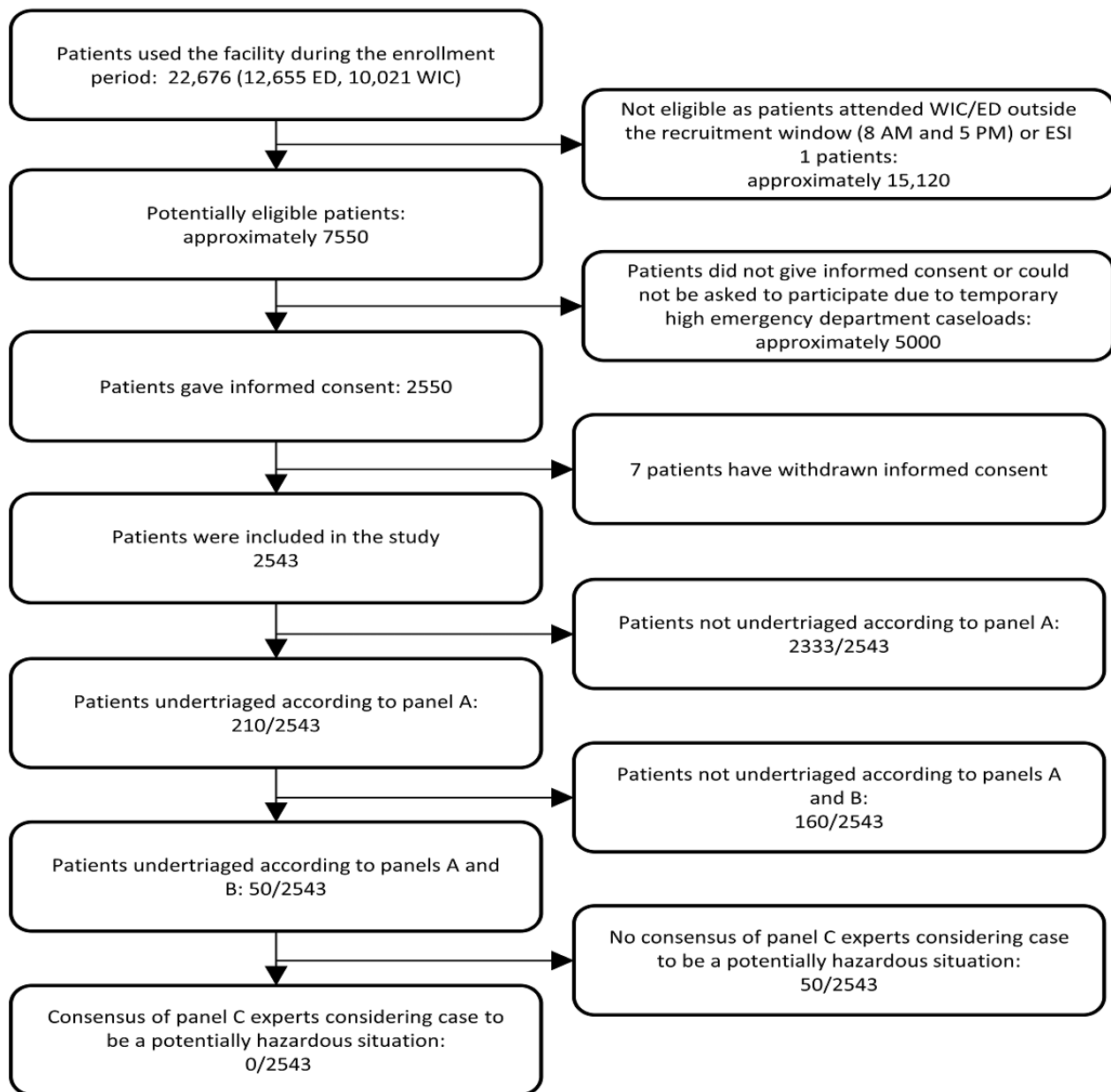
Table 3. Characteristics of the study population (N=2543).

Characteristics	Participants, n (%)
Age (years)	
18-49	1397 (54.94)
50-65	668 (26.27)
66-80	360 (14.16)
>80	118 (4.64)
Gender	
Female	1227 (48.25)
Male	1316 (51.75)
Reason for encounter (15 most frequent)	
Stomach pain	287 (11.26)
Chest pain	168 (6.61)
Lumbar back pain	144 (5.66)
Urinary tract problems	124 (4.88)
Trauma or fall	121 (4.76)
Headache	90 (3.54)
Dizziness	87 (3.42)
Wound or skin injury	82 (3.22)
Foot injury (caused by an accident)	81 (3.19)
Leg problems	74 (2.91)
Breathlessness	69 (2.71)
Cold or influenza infection	64 (2.52)
Finger injury (caused by an accident)	55 (2.16)
Knee injury (caused by an accident)	51 (2.01)
Hand injury (caused by an accident)	43 (1.69)

Table 4. Distribution of cases according to the various triage levels, as defined in (N=2543).

Triage level	Participants, n (%)
Emergency ambulance	57 (0.02)
Emergency hospital	142 (0.06)
Immediately ambulance	2 (0)
Immediately hospital	685 (0.27)
Immediately doctor	844 (0.33)
Today hospital	3 (0)
Today doctor	579 (0.23)
Later doctor	36 (0.01)
Today call center	26 (0.01)
Later call center	60 (0.02)
Today pharmacy	0 (0)
Later pharmacy	30 (0.01)
Today self-care	0 (0)
Later self-care	77 (0.03)
Unclear	2 (0)

Figure 1. Flow of patients through this study and triage assessment steps by expert panels A, B, and C. ED: emergency department; ESI: Emergency Severity Index; WIC: walk-in clinic.



In 210 (8.26%) of the 2543 cases, the recommendation issued by the symptom-checker was below the range of appropriate triage levels defined by the panel A experts and therefore undertriaged. Further, 50 (1.96%) of these 210 patients were equally undertriaged according to panel B. However, for none of these 50 patients did panel C reach a consensus that the undertriage was potentially hazardous. This resulted in an upper 95% confidence bound for the probability of a potentially hazardous undertriage of 0.1184%. If the criterion for potentially

hazardous undertriage was defined as a majority of panel C members considering life-threatening or harmful self-triage “rather likely” or “likely,” 4 of the 50 cases fulfilled this criterion. This resulted in an upper 95% confidence bound for the probability of a potentially hazardous undertriage of 0.3616%.

Table 5 shows the adjudication of potentially hazardous undertriage for all 50 cases evaluated by the experts of panel C.

Table 5. Distribution of assessment for potentially hazardous undertriage for all 50 cases, as adjudicated by each of the 5 members of panel C.

Case number	Unlikely	Rather unlikely	Rather likely	Likely
Case 1	2	3	0	0
Case 2	5	0	0	0
Case 3	3	2	0	0
Case 4	3	2	0	0
Case 5	4	0	1	0
Case 6	5	0	0	0
Case 7	2	2	1	0
Case 8	4	1	0	0
Case 9	0	2	2	1
Case 10	4	0	1	0
Case 11	5	0	0	0
Case 12	5	0	0	0
Case 13	1	1	2	1
Case 14	4	1	0	0
Case 15	2	2	1	0
Case 16	5	0	0	0
Case 17	4	1	0	0
Case 18	5	0	0	0
Case 19	5	0	0	0
Case 20	5	0	0	0
Case 21	3	1	1	0
Case 22	2	1	1	1
Case 23	3	1	1	0
Case 24	5	0	0	0
Case 25	5	0	0	0
Case 26	5	0	0	0
Case 27	5	0	0	0
Case 28	5	0	0	0
Case 29	2	1	2	0
Case 30	4	0	1	0
Case 31	5	0	0	0
Case 32	2	1	1	1
Case 33	4	1	0	0
Case 34	4	1	0	0
Case 35	5	0	0	0
Case 36	0	2	3	0
Case 37	4	1	0	0
Case 38	5	0	0	0
Case 39	5	0	0	0
Case 40	2	1	1	1
Case 41	1	3	1	0
Case 42	5	0	0	0

Case number	Unlikely	Rather unlikely	Rather likely	Likely
Case 43	5	0	0	0
Case 44	5	0	0	0
Case 45	5	0	0	0
Case 46	5	0	0	0
Case 47	5	0	0	0
Case 48	1	0	4	0
Case 49	4	1	0	0
Case 50	5	0	0	0

The central (2-sided) 95% Clopper-Pearson CI for the probability of undertriage according to panel A is 7.22% to 9.40%. The central (2-sided) 95% Clopper-Pearson CI for the probability of overtriage according to panel A (450 cases, 17.69%) is 16.23% to 19.24%.

For the 50 out of 2543 cases that were undertriaged according to the judgments of panels A and B, the central (2-sided) 95% Clopper-Pearson CI for the corresponding probability is 1.539% to 2.688%.

The central (2-sided) 95% Clopper-Pearson CI for the probability of a potentially hazardous undertriage for the consensus criterion (0 out of 2543 cases) is 0% to 0.1458% and 0.0431% to 0.4045%, according to the majority criterion (4 out of 2543 cases).

Discussion

Principal Findings

Our study corroborates the safety of the SMASS pathfinder symptom-checker for medical self-assessment of acute complaints in a real-life clinical setting. A stepwise evaluation of 2543 consecutive patients by 3 independent expert panels yielded no cases of potentially hazardous undertriage when the consensus criterion was applied and 4 cases when the majority criterion was applied.

In a systematic literature search, we found insufficient evidence from comparatively small studies for the safe use of symptom-checkers in clinical routine (Demurtas et al, unpublished data, 2021). Further, 1 study with 825 patients showing “exactly matched” triage in 52.6% has been published in abstract form only [15]. Another study yielded correct triage in only 50%-74% of cases [16]. A third study, from Germany, evaluated the safety of urgency advice provided to 378 patients at an interdisciplinary ED by a symptom-checker [17], showing undertriage in 34 (8.9%) and overtriage in 216 (57.1%) cases. A potentially hazardous situation was identified in 20 (5.3%) cases. This figure appears considerably higher than our finding, although an interrater variability was not taken into account in the German study. Another study aimed to analyze the performance of a clinical decision support system that allowed patients to self-triage in the ED of a university hospital. The authors concluded that the self-triage device was safe, as the assessments by the system and the physicians were congruent concerning the classification as an emergency. However, in

contrast to our study, the risk to life or health was not assessed [18].

In the absence of a broad study base, we cannot compare our results with previous, similarly designed studies for symptom-checkers. In contrast, medical telephone triage has been extensively evaluated during the last 25 years [19-24] and has gained broad clinical support, despite ambivalent conclusions regarding safety.

In a systematic review analyzing 13 observational studies and 10 studies that simulated high-risk patients, safe triage was found to be 46% to 97% [25]. Another systematic review involving computer-assisted telephone triage in urgent care [26] pointed out 4 studies that indicated potential undertriage errors [27-30]. Notably, hospitalization rates of patients who were advised to seek nonurgent care ranged from 9.2% to 48%. Potentially life-threatening situations emerged in 0.84% of cases according to 1 study [29].

We have previously investigated the safety of computer-assisted telephone triage in 208 patients with non-life threatening conditions consulting the ED at a university hospital [31]. We found poor agreement between the assessments by the call center, the emergency physician, and the general practitioners who later cared for the patients. In 1 case, a risk to health or life was found.

The Cochrane Collaboration in their 2004 systematic review on telephone triage concluded that insufficient data existed regarding safety [32]. In light of the available information, the results of our study compare favorably to the published data on telephone triage.

Our study has several strengths and weaknesses. We included a large number of patients in a real-world clinical setting. In addition, this study's design enabled us to eliminate the low interrater reliability of medical triage decisions by having 3 independent expert panels. This allows robust conclusions about the safety of the evaluated symptom-checker.

For reasons of feasibility, we performed our study in a hospital setting, where patients were triaged to the WIC or ED according to ESI criteria. Thus, a wide variety of cases could be assessed. On the other hand, the symptom-checker was not used in a setting outside the hospital, limiting generalizability. However, presenting symptoms largely overlap with those encountered in primary care, and a potential selection bias toward more

severe cases would support the conclusion on the device's safety if it were used in primary care.

A potential limitation of our study is its single-center design. However, the Cantonal Hospital Baden serves a mixed urban and rural population of approximately 300,000 people and offers all medical services except cardiac surgery and neurosurgery. We therefore believe that the patient sample in our study is fairly representative of the general population.

The total number of patients frequenting the WIC and ED during the time of recruitment was 22,676; thus, only approximately 11% of them participated in this study, potentially resulting in selection bias. Due to limited resources, inclusions were possible only during the daytime, leaving approximately 7550 potential participants. Further, 1.5% (340/22,676) were ESI 1 patients, who were not eligible for this study. It could be speculated that patients visiting an ED at night time might be more seriously ill than those during the daytime. This potential bias would make our cohort more comparable to a setting in primary care.

In our study, we have focused on the safety of the symptom-checker. A possible limitation may have resulted from the fact that each case was initially assessed by a single member of panel A. This could have precluded passing a potentially hazardous case to panel B. While maximum patient safety may theoretically be desirable, it should be weighed against the disadvantages of overtriage, notably inefficiency, unnecessary referrals, and a higher risk of overmedicalization, all of which increase costs. In our study, the overtriage rate after assessment

by panel A was 17.69% (450 cases). This figure is comparable to published rates of overtriage by teleconsultation and teletriage, which range from 12% to 57% [33-37]. In a further round of data analysis, we will also have the overtriaged cases assessed by panel B to include the low interrater reliability in the analysis. As with undertriaged cases, this is likely to reduce the overtriaged cases.

From the end of February, the COVID-19 pandemic required special hygiene measures for the tablet computers used, making patient recruitment more difficult as the first wave of the pandemic peaked in March 2020. The pandemic is also likely to have affected the case mix, which may have shifted slightly toward COVID-19-positive patients.

The urgency grading used in the 2D matrix for the triage levels (Table 1) was defined at the discretion of this study's team, implicating a certain degree of subjectiveness. While the range of appropriate triage levels was defined based on this order, the experts did not always explicitly mark all of the intermediate triage levels as appropriate.

Conclusions

The SMASS pathfinder symptom-checker proved to be a safe triage tool, avoiding undertriage in a real-life clinical setting of emergency consultations at a walk-in clinic and ED. Although for practical reasons the symptom-checker was not evaluated outside the hospital environment, our data do not suggest that its safety may have been compromised if used for self-triage by patients in a domestic setting.

Acknowledgments

We thank this study's nurses of the clinical trial unit, Simone Fontana, Stefanie Leuenberger, and Franziska Rutz, PhD, for their excellent work as well as Elena Righi, PhD, and Lee Smith, PhD, for proofreading this paper. This study was funded by the Health Innovation Hub of the Cantonal Hospital Baden, Switzerland. The symptom-checker was provided by in4medicine, Inc, at no charge.

Data Availability

The raw data of our study is available from in4medicine, Inc [38].

Authors' Contributions

AM, JR, MS, and MV were responsible for the concept and design. PR, MS, and BG acquired the data, which were analyzed and interpreted by MV, AM, JD, and JR. The first draft was written by AM and then critically revised in conjunction with JR, JD, and MV. Data analysis was done by MV, with oversight by AM and JR.

Conflicts of Interest

AM is the founder and chief executive officer of in4medicine, Inc. BG is a part-time employee of in4medicine, Inc, and JD received a scientific grant from in4medicine, Inc. MV received an honorarium from in4medicine, Inc. The other authors report no conflicts of interest.

References

1. Thierrin C, Augsburger A, Dami F, Monney C, Staeger P, Clair C. Impact of a telephone triage service for non-critical emergencies in Switzerland: a cross-sectional study. *PLoS One*. 2021;16(4):1-13. [FREE Full text] [doi: [10.1371/journal.pone.0249287](https://doi.org/10.1371/journal.pone.0249287)] [Medline: [33798216](https://pubmed.ncbi.nlm.nih.gov/33798216/)]
2. Steeman L, Uijen M, Plat E, Huibers L, Smits M, Giesen P. Out-of-hours primary care in 26 European countries: an overview of organizational models. *Fam Pract*. 2020;37(6):744-750. [FREE Full text] [doi: [10.1093/fampra/cmaa064](https://doi.org/10.1093/fampra/cmaa064)] [Medline: [32597962](https://pubmed.ncbi.nlm.nih.gov/32597962/)]

3. Regulation (EU) 2017/745 of The European Parliament and of the Council of 5 April 2017 on medical devices, amending Directive 2001/83/EC, Regulation (EC) No 178/2002 and Regulation (EC) No 1223/2009 and repealing Council Directives 90/385/EEC and 93/42/EEC. Official Journal of the European Union. URL: <https://eur-lex.europa.eu/legal-content/EN/TXT/HTML/?uri=CELEX:32017R0745> [accessed 2024-05-31]
4. Fraser H, Coiera E, Wong D. Safety of patient-facing digital symptom checkers. *Lancet*. 2018;392(10161):2263-2264. [doi: [10.1016/S0140-6736\(18\)32819-8](https://doi.org/10.1016/S0140-6736(18)32819-8)] [Medline: [30413281](https://pubmed.ncbi.nlm.nih.gov/30413281/)]
5. 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*. 2023;25:e43803. [FREE Full text] [doi: [10.2196/43803](https://doi.org/10.2196/43803)] [Medline: [37266983](https://pubmed.ncbi.nlm.nih.gov/37266983/)]
6. El-Osta A, Webber I, Alaa A, Bagkeris E, Mian S, Taghavi Azar Sharabiani M, et al. What is the suitability of clinical vignettes in benchmarking the performance of online symptom checkers? An audit study. *BMJ Open*. 2022;12(4):e053566. [FREE Full text] [doi: [10.1136/bmjopen-2021-053566](https://doi.org/10.1136/bmjopen-2021-053566)] [Medline: [35477872](https://pubmed.ncbi.nlm.nih.gov/35477872/)]
7. Munro J, Clancy M, Knowles E, Sampson F, Nicholl J. Evaluation of NHS direct: impact and appropriateness: final report of the phase 1 research. *NHS Direct*. 2001. URL: https://www.researchgate.net/publication/246112002_Evaluation_of_NHS_Direct_first_wave_sites_Final_report_of_the_phase_1_research [accessed 2024-05-02]
8. Snooks H, Peconi J, Munro J, Cheung WY, Rance J, Williams A. An evaluation of the appropriateness of advice and healthcare contacts made following calls to NHS direct Wales. *BMC Health Serv Res*. 2009;9:178. [FREE Full text] [doi: [10.1186/1472-6963-9-178](https://doi.org/10.1186/1472-6963-9-178)] [Medline: [19793398](https://pubmed.ncbi.nlm.nih.gov/19793398/)]
9. Semigran HL, Linder JA, Gidengil C, Mehrotra A. Evaluation of symptom checkers for self diagnosis and triage: audit study. *BMJ*. 2015;351:h3480. [FREE Full text] [doi: [10.1136/bmj.h3480](https://doi.org/10.1136/bmj.h3480)] [Medline: [26157077](https://pubmed.ncbi.nlm.nih.gov/26157077/)]
10. Chambers D, Cantrell A, Johnson M, Preston L, Baxter S, 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. 2019;7(29):1-87. [doi: [10.3310/hsdr07290](https://doi.org/10.3310/hsdr07290)] [Medline: [31433612](https://pubmed.ncbi.nlm.nih.gov/31433612/)]
11. Christ M, Bingisser R, Nickel CH. Emergency Triage. An Overview. *Dtsch Med Wochenschr*. 2016;141(5):329-335. [doi: [10.1055/s-0041-109126](https://doi.org/10.1055/s-0041-109126)] [Medline: [26939102](https://pubmed.ncbi.nlm.nih.gov/26939102/)]
12. St George I, Cullen M, Branney M. Healthline: do primary care doctors agree with the advice? *N Z Med J*. 2005;118(1224):U1693. [Medline: [16258573](https://pubmed.ncbi.nlm.nih.gov/16258573/)]
13. Gribben B. General practitioners' assessments of the primary care caseload in Middlemore Hospital Emergency Department. *N Z Med J*. 2003;116(1169):U329. [Medline: [12601406](https://pubmed.ncbi.nlm.nih.gov/12601406/)]
14. Durand AC, Gentile S, Gerbeaux P, Alazia M, Kiegel P, Luigi S, et al. Be careful with triage in emergency departments: interobserver agreement on 1,578 patients in France. *BMC Emerg Med*. 2011;11:19. [FREE Full text] [doi: [10.1186/1471-227X-11-19](https://doi.org/10.1186/1471-227X-11-19)] [Medline: [22040017](https://pubmed.ncbi.nlm.nih.gov/22040017/)]
15. Koskela T, Liu V, Kaila M. How does triage by an electronic symptom checker match with triage by a nurse? *Stud Health Technol Inform*. 2022;294:571-572. [doi: [10.3233/SHTI220528](https://doi.org/10.3233/SHTI220528)] [Medline: [35612149](https://pubmed.ncbi.nlm.nih.gov/35612149/)]
16. Yu SWY, Ma A, Tsang VHM, Chung LSW, Leung SC, Leung LP. Triage accuracy of online symptom checkers for accident and emergency department patients. *Hong Kong J Emerg Med*. 2020;27(4):217-222. [doi: [10.1177/1024907919842486](https://doi.org/10.1177/1024907919842486)]
17. 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*. 2022;10(3):e32340. [FREE Full text] [doi: [10.2196/32340](https://doi.org/10.2196/32340)] [Medline: [35343909](https://pubmed.ncbi.nlm.nih.gov/35343909/)]
18. Määttä J, Lindell R, Hayward N, Martikainen S, Honkanen K, Inkala M, et al. Diagnostic performance, triage safety, and usability of a clinical decision support system within a university Hospital emergency department: algorithm performance and usability study. *JMIR Med Inform*. 2023;11:e46760. [FREE Full text] [doi: [10.2196/46760](https://doi.org/10.2196/46760)] [Medline: [37656018](https://pubmed.ncbi.nlm.nih.gov/37656018/)]
19. Derckx HP, Rethans JJE, Muijtjens AM, Maiburg BH, Winkens R, van Rooij HG, et al. Quality of clinical aspects of call handling at Dutch out of hours centres: cross sectional national study. *BMJ*. 2008;337:a1264. [FREE Full text] [doi: [10.1136/bmj.a1264](https://doi.org/10.1136/bmj.a1264)] [Medline: [18790814](https://pubmed.ncbi.nlm.nih.gov/18790814/)]
20. Derckx HP, Rethans JJE, Maiburg BH, Winkens RA, Muijtjens AM, van Rooij HG, et al. Quality of communication during telephone triage at Dutch out-of-hours centres. *Patient Educ Couns*. 2009;74(2):174-178. [doi: [10.1016/j.pec.2008.08.002](https://doi.org/10.1016/j.pec.2008.08.002)] [Medline: [18845413](https://pubmed.ncbi.nlm.nih.gov/18845413/)]
21. Niemann S, Meer A, Simonin C, Abel T. Medical telephone triage and patient behaviour: how do they compare? *Swiss Med Wkly*. 2004;134(9-10):126-131. [doi: [10.4414/smw.2004.10276](https://doi.org/10.4414/smw.2004.10276)] [Medline: [15106022](https://pubmed.ncbi.nlm.nih.gov/15106022/)]
22. Lattimer V, George S, Thompson F, Thomas E, Mullee M, Turnbull J, et al. Safety and effectiveness of nurse telephone consultation in out of hours primary care: randomised controlled trial. The South Wiltshire Out of Hours Project (SWOOP) Group. *BMJ*. 1998;317(7165):1054-1059. [doi: [10.1136/bmj.317.7165.1054](https://doi.org/10.1136/bmj.317.7165.1054)] [Medline: [9774295](https://pubmed.ncbi.nlm.nih.gov/9774295/)]
23. Campbell JL, Fletcher E, Britten N, Green C, Holt TA, Lattimer V, et al. Telephone triage for management of same-day consultation requests in general practice (the ESTEEM trial): a cluster-randomised controlled trial and cost-consequence analysis. *Lancet*. 2014;384(9957):1859-1868. [FREE Full text] [doi: [10.1016/S0140-6736\(14\)61058-8](https://doi.org/10.1016/S0140-6736(14)61058-8)] [Medline: [25098487](https://pubmed.ncbi.nlm.nih.gov/25098487/)]
24. Murdoch J, Varley A, Fletcher E, Britten N, Price L, Calitri R, et al. Implementing telephone triage in general practice: a process evaluation of a cluster randomised controlled trial. *BMC Fam Pract*. 2015;16:47. [FREE Full text] [doi: [10.1186/s12875-015-0263-4](https://doi.org/10.1186/s12875-015-0263-4)] [Medline: [25887747](https://pubmed.ncbi.nlm.nih.gov/25887747/)]

25. Huibers L, Smits M, Renaud V, Giesen P, Wensing M. Safety of telephone triage in out-of-hours care: a systematic review. *Scand J Prim Health Care*. 2011;29(4):198-209. [FREE Full text] [doi: [10.3109/02813432.2011.629150](https://doi.org/10.3109/02813432.2011.629150)] [Medline: [22126218](https://pubmed.ncbi.nlm.nih.gov/22126218/)]
26. Sexton V, Dale J, Bryce C, Barry J, Sellers E, Atherton H. Service use, clinical outcomes and user experience associated with urgent care services that use telephone-based digital triage: a systematic review. *BMJ Open*. 2022;12(1):e051569. [FREE Full text] [doi: [10.1136/bmjopen-2021-051569](https://doi.org/10.1136/bmjopen-2021-051569)] [Medline: [34980613](https://pubmed.ncbi.nlm.nih.gov/34980613/)]
27. Foster J, Jessopp L, Chakraborti S. Do callers to NHS direct follow the advice to attend an accident and emergency department? *Emerg Med J*. 2003;20(3):285-288. [FREE Full text] [doi: [10.1136/emj.20.3.285](https://doi.org/10.1136/emj.20.3.285)] [Medline: [12748156](https://pubmed.ncbi.nlm.nih.gov/12748156/)]
28. Sprivilis P, Carey M, Rouse I. Compliance with advice and appropriateness of emergency presentation following contact with the HealthDirect telephone triage service. *Emerg Med Australas*. 2004;16(1):35-40. [doi: [10.1111/j.1742-6723.2004.00538.x](https://doi.org/10.1111/j.1742-6723.2004.00538.x)] [Medline: [15239753](https://pubmed.ncbi.nlm.nih.gov/15239753/)]
29. Dale J, Higgins J, Williams S, Foster T, Snooks H, Crouch R, et al. Computer assisted assessment and advice for "non-serious" 999 ambulance service callers: the potential impact on ambulance despatch. *Emerg Med J*. 2003;20(2):178-183. [doi: [10.1136/emj.20.2.178](https://doi.org/10.1136/emj.20.2.178)] [Medline: [12642540](https://pubmed.ncbi.nlm.nih.gov/12642540/)]
30. Stewart B, Fairhurst R, Markland J, Marzouk O. Review of calls to NHS direct related to attendance in the paediatric emergency department. *Emerg Med J*. 2006;23(12):911-914. [FREE Full text] [doi: [10.1136/emj.2006.039339](https://doi.org/10.1136/emj.2006.039339)] [Medline: [17130596](https://pubmed.ncbi.nlm.nih.gov/17130596/)]
31. Meer A, Gwerder T, Duembgen L, Zumbrennen N, Zimmermann H. Is computer-assisted telephone triage safe? A prospective surveillance study in walk-in patients with non-life-threatening medical conditions. *Emerg Med J*. 2012;29(2):124-128. [doi: [10.1136/emj.2009.080614](https://doi.org/10.1136/emj.2009.080614)] [Medline: [20961939](https://pubmed.ncbi.nlm.nih.gov/20961939/)]
32. Bunn F, Byrne G, Kendall S. Telephone consultation and triage: effects on health care use and patient satisfaction. *Cochrane Database Syst Rev*. 2004;(4):CD004180. [doi: [10.1002/14651858.CD004180.pub2](https://doi.org/10.1002/14651858.CD004180.pub2)] [Medline: [15495083](https://pubmed.ncbi.nlm.nih.gov/15495083/)]
33. Brasseur E, Gilbert A, Donneau AF, Monseur J, Ghuysen A, D'Orio V. Reliability and validity of an original nurse telephone triage tool for out-of-hours primary care calls: the SALOMON algorithm. *Acta Clin Belg*. 2022;77(3):640-646. [doi: [10.1080/17843286.2021.1936353](https://doi.org/10.1080/17843286.2021.1936353)] [Medline: [34081571](https://pubmed.ncbi.nlm.nih.gov/34081571/)]
34. Nørøxe KB, Huibers L, Moth G, Vedsted P. Medical appropriateness of adult calls to Danish out-of-hours primary care: a questionnaire-based survey. *BMC Fam Pract*. 2017;18(1):34. [FREE Full text] [doi: [10.1186/s12875-017-0617-1](https://doi.org/10.1186/s12875-017-0617-1)] [Medline: [28292257](https://pubmed.ncbi.nlm.nih.gov/28292257/)]
35. Cook R, Thakore S, Morrison W, Meikle J. To ED or not to ED: NHS 24 referrals to the emergency department. *Emerg Med J*. 2010;27(3):213-215. [doi: [10.1136/emj.2008.064261](https://doi.org/10.1136/emj.2008.064261)] [Medline: [20304891](https://pubmed.ncbi.nlm.nih.gov/20304891/)]
36. Giesen P, Ferwerda R, Tijssen R, Mookink H, Drijver R, van den Bosch W, et al. Safety of telephone triage in general practitioner cooperatives: do triage nurses correctly estimate urgency? *Qual Saf Health Care*. 2007;16(3):181-184. [FREE Full text] [doi: [10.1136/qshc.2006.018846](https://doi.org/10.1136/qshc.2006.018846)] [Medline: [17545343](https://pubmed.ncbi.nlm.nih.gov/17545343/)]
37. Scarfone RJ, Luberti AA, Mistry RD. Outcomes of children referred to an emergency department by an after-hours call center. *Pediatr Emerg Care*. 2004;20(6):367-372. [doi: [10.1097/01.pec.0000133610.42699.41](https://doi.org/10.1097/01.pec.0000133610.42699.41)] [Medline: [15179144](https://pubmed.ncbi.nlm.nih.gov/15179144/)]
38. 2024_05_01_Data_ms_58157. in4medicine. URL: https://pub.in4medicine.ch/fileadmin/pub/2024_05_01_Data_ms_58157.pdf [accessed 2024-06-01]

Abbreviations

- ED:** emergency department
- ESI:** Emergency Severity Index
- SMASS:** Swiss Medical Assessment System
- WIC:** walk-in clinic

Edited by A Mavragani; submitted 07.03.24; peer-reviewed by C Lowe, R Payne; comments to author 30.03.24; revised version received 15.04.24; accepted 29.05.24; published 27.06.24

Please cite as:

Meer A, Rahm P, Schwendinger M, Vock M, Grunder B, Demurtas J, Rutishauser J
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
URL: <https://www.jmir.org/2024/1/e58157>
doi: [10.2196/58157](https://doi.org/10.2196/58157)
PMID: [38809606](https://pubmed.ncbi.nlm.nih.gov/38809606/)

©Andreas Meer, Philipp Rahm, Markus Schwendinger, Michael Vock, Bettina Grunder, Jacopo Demurtas, Jonas Rutishauser. Originally published in the Journal of Medical Internet Research (<https://www.jmir.org>), 27.06.2024. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in the Journal of Medical Internet Research (ISSN 1438-8871), is properly cited. The complete bibliographic information, a link to the original publication on <https://www.jmir.org/>, as well as this copyright and license information must be included.