Digital self-triage: an evaluation of patient compliance

Anthony Pairon, University of Antwerp

Promotor: prof. dr. Hilde Philips, University of Antwerp

Co-promotoren: prof. dr. Veronique Verhoeven, University of Antwerp

Master of Family Medicine
Masterproef
Academiejaar: 2022 – 2023

Huisartsgeneeskunde

Deze masterproef is een examendocument dat niet werd gecorrigeerd voor eventueel vastgestelde fouten. Zonder voorafgaande schriftelijke toestemming van zowel de promotor(en) als de auteur(s) is overnemen, kopiëren, gebruiken of realiseren van deze uitgave of gedeelten ervan verboden. Voor aanvragen tot of informatie i.v.m. het overnemen en/of gebruik en/of realisatie van gedeelten uit deze publicatie, wendt u tot de universiteit waaraan de auteur is ingeschreven.

Voorafgaande schriftelijke toestemming van de promotor(en) is eveneens vereist voor het aanwenden van de in dit afstudeerwerk beschreven (originele) methoden, producten, schakelingen en programma's voor industrieel of commercieel nut en voor de inzending van deze publicatie ter deelname aan wetenschappelijke prijzen of wedstrijden.

Abstract (Nederlands)

Achtergrond

Online symptoomcheckers en digitale triagesystemen maken deel uit van de huidige inspanningen om patiënten betere bijstand te bieden in hun hulpzoekgedrag. Het huidige bewijs over deze systemen vertoont een aanzienlijke variatie in onderzoeksopzet en resultaten. Bovendien zijn veel elementen nog niet onderzocht. Vooral de therapietrouw vereist een meer wetenschappelijk onderbouwde aanpak. Het huidige, beperkte bewijs is gericht op de intentie om het advies na te volgen, in plaats van op de kwantificering van de directe impact op de besluitvorming van patiënten.

Onderzoeksvragen

- RQ1: In welke mate gaan mensen in op het advies van een digitale zelf-triagetool?
    - RQ1.1: Hoe beïnvloeden patiënt-, urgentie-, ziekte- en triage-kenmerken de therapietrouw?
- RQ2: In hoeverre hangen intentie tot naleving en naleving samen?

Methode

Een prospectieve longitudinale observationele studie werd ontworpen, waarbij de online triagetool moetiknaardedokter.be werd geïntegreerd in een GPC-website. Routinematige registratiegegevens van het triagesysteem werden gekoppeld aan een tweede dataset van ICAREdata, met anonieme patiëntgegevens van de niet-planbare zorgsystemen. Algoritmische data-matching visualiseerde de patiëntenstroom door het zorgsysteem, waardoor de naleving van het gegeven advies door de patiënt kon worden gekwantificeerd. (n = 398) Een gestructureerde vragenlijst, voorgelegd aan het einde van de triageprocedure, peilde de intentie van de patiënt vóór en na het advies, de impact van de triage op het hulpzoekgedrag, en de onderliggende motivaties. (n = 99) Deze resultaten werden vergeleken met de observationele dataset, waarbij de intentie tot navolging met de effectieve navolging werd vergeleken.

Resultaten

Over het algemeen voldeed 67,3% aan het gegeven advies, wat steeg tot 77% voor degenen die een niet-dringend triageadvies kregen (U5). Naleving was gerelateerd aan gebruik in het weekend versus doordeweekse nachten (OR 2,9 95% CI 1,3 - 3,9, p<0,002). 97% van de respondenten meende een arts nodig te hebben voorafgaand aan de triage. Na gebruik gaf 88,9% aan dat zij van plan waren zich hieraan te houden, maar in 26,8% van de geobserveerde gevallen week het gedrag af van de verklaarde intentie.

Conclusie

De naleving van het triageadvies door patiënten is over het algemeen hoog voor moetiknaardedokter.be, hetgeen overeenstemt met de huidige kennis. Er is een aanzienlijke discrepantie tussen het gedrag van patiënten en hun verklaarde intentie, waardoor voorzichtigheid geboden is bij het gebruik van dit laatste als proximale parameter. Toekomstig onderzoek naar naleving moet van relevante omvang en reikwijdte zijn, en uitgevoerd worden binnen de beoogde setting van de systemen.

Abstract (English)

Background

Online symptom checkers and digital triage tools are part of a current effort to better assist patients in their healthcare seeking behavior. Present evidence on these systems presents a considerable variation in study design and results, with many elements remaining unexplored. Compliance in particular requires a more solid evidence-based approach, with limited evidence currently focusing on intention to comply, rather than quantifying the direct impact on patients' decision-making.

Research questions

- RQ1: How compliant are people with the advice given by a digital self-triage tool?
- RQ1.1: How do patient, urgency, disease, and triage characteristics influence compliance?
- RQ2: To what degree do intention to comply and compliance correlate?

Method

A prospective longitudinal observational study was designed in which the online triage tool moetiknaardedokter.be was integrated into a GPC website. Routine registration data of the triage system was linked to a second dataset from ICAREdata, containing anonymous patient data from unplannable care systems. Algorithmic data-matching visualized the patient flow through the healthcare system, enabling the quantification of patient compliance to the advice given. (n = 398)
A structured questionnaire, presented at the end of the triage procedure, gaged the patient's intention pre- and post-advice, the impact of triage on care seeking behavior, and underlying motivational drivers. (n = 99) These results were matched to the observational dataset, comparing intention to subsequent action.

Results

Overall, 67.3% complied to the advice given, which increased to 77% for those receiving a non-urgent triage advice (U5). Compliance was associated to utilization in the weekend vs. weeknight (OR 2.9 95% CI 1.3 – 3.9, $p<0.002$).97% of survey respondents believed they required a doctor prior to triage. After utilization, 88.9% indicated they intended to comply, yet in 26.8% of cases observed, behavior differed from their stated intention.

Conclusion

Patient adherence to triage advice is generally high for moetiknaardedokter.be and in line with current evidence. There is considerable discrepancy between patient behavior and their declared intention, warranting caution when using the latter as a proximal outcome. Future research on compliance should be of relevant scale and scope and be performed within the tools' intended setting.

Table of contents

Study 1: A scoping review on the use and usefulness of online symptom checkers and triage systems: How to proceed?

1. Abstract....01
2. Introduction....02
   2.1. Objective
3. Methods....02
   3.1. Sources and search strategy
   3.2. Selection criteria
   3.3. Study selection
   3.4. Data charting and data extraction
4. Results....03
   4.1. Study outcomes
      4.1.1. Demographics
      4.1.2. Safety
      4.1.3. Accuracy
         4.1.3.1. Diagnostic accuracy
         4.1.3.2. Triage accuracy
      4.1.4. Compliance
      4.1.5. Impact
         4.1.5.1. Impact on health literacy
         4.1.5.2. Impact on health equity and digital divide
         4.1.5.3. Impact on OOH-care
            4.1.5.3.1. Healthcare seeking behavior
            4.1.5.3.2. Workload
         4.1.5.4. Accessibility to care
         4.1.5.5. Impact on public health
      4.1.6. Cost effectiveness
      4.1.7. User experience
         4.1.7.1. Patients
         4.1.7.2. Caregivers
      4.1.8. Combination of demand management measures
5. Discussion....09
   5.1. Limitations on this review
6. Conclusion....11

Author contributions
Funding
Conflict of interest
Publisher's note
Supplementary material
References....12

Study 2: Digital self-triage: An evaluation of patient compliance

1. Introduction....13
   1.1. Objective
2. Methods....13
   2.1. Study design and setting
   2.2. The tool
   2.3. Data
   2.4. Outcomes
   2.5. Statistical analysis

3. Results....15
3.1. Participants
3.2. User demographics
3.3. Triage characteristics
3.4. Compliance
4. Discussion....17
4.1. User demographics
4.2. Triage characteristics
4.3. Compliance
4.4. Study limitations
4.5. Future research
References....19
Appendix 1. Approval of the ethics committee of UZA-UAntwerp
Appendix 2. Approved study protocol

frontiers | Frontiers in Medicine

TYPE Review
PUBLISHED 06 January 2023
DOI 10.3389/fmed.2022.1040926

Check for updates

OPEN ACCESS

EDITED BY
Marija Petek Šter,
University of Ljubljana, Slovenia

REVIEWED BY
Janet Michel,
University Hospital of Bern, Switzerland
Sundus Tariq,
The University of Faisalabad, Pakistan

*CORRESPONDENCE
Anthony Pairon
✉ anthony.pairon@student.uantwerpen.be

SPECIALTY SECTION
This article was submitted to Family Medicine and Primary Care, a section of the journal Frontiers in Medicine

RECEIVED 09 September 2022
ACCEPTED 16 December 2022
PUBLISHED 06 January 2023

CITATION
Pairon A, Philips H and Verhoeven V (2023) A scoping review on the use and usefulness of online symptom checkers and triage systems: How to proceed?
Front. Med. 9:1040926.
doi: 10.3389/fmed.2022.1040926

COPYRIGHT
© 2023 Pairon, Philips and Verhoeven. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

A scoping review on the use and usefulness of online symptom checkers and triage systems: How to proceed?

Anthony Pairon*, Hilde Philips and Veronique Verhoeven

Department of Family Medicine and Population Health, University of Antwerp, Antwerp, Belgium

Background: Patients are increasingly turning to the Internet for health information. Numerous online symptom checkers and digital triage tools are currently available to the general public in an effort to meet this need, simultaneously acting as a demand management strategy to aid the overburdened health care system. The implementation of these services requires an evidence-based approach, warranting a review of the available literature on this rapidly evolving topic.

Objective: This scoping review aims to provide an overview of the current state of the art and identify research gaps through an analysis of the strengths and weaknesses of the presently available literature.

Methods: A systematic search strategy was formed and applied to six databases: Cochrane library, NICE, DARE, NIHR, Pubmed, and Web of Science. Data extraction was performed by two researchers according to a pre-established data charting methodology allowing for a thematic analysis of the results.

Results: A total of 10,250 articles were identified, and 28 publications were found eligible for inclusion. Users of these tools are often younger, female, more highly educated and technologically literate, potentially impacting digital divide and health equity. Triage algorithms remain risk-averse, which causes challenges for their accuracy. Recent evolutions in algorithms have varying degrees of success. Results on impact are highly variable, with potential effects on demand, accessibility of care, health literacy and syndromic surveillance. Both patients and healthcare providers are generally positive about the technology and seem amenable to the advice given, but there are still improvements to be made toward a more patient-centered approach. The significant heterogeneity across studies and triage systems remains the primary challenge for the field, limiting transferability of findings.

Conclusion: Current evidence included in this review is characterized by significant variability in study design and outcomes, highlighting the significant challenges for future research.

An evolution toward more homogeneous methodologies, studies tailored to the intended setting, regulation and standardization of evaluations, and a patient-centered approach could benefit the field.

KEYWORDS

triage, symptom checkers, diagnosis, digital health, mhealth (mobile health), ehealth

1. Introduction

Out-of-hours (OOH) medical care is currently facing an increasing demand, affecting both general practitioner cooperatives (GPCs) and emergency departments (ED) (1). This rise in workload has multifactorial origins and can be partially explained by macro-evolutions such as a population that is both growing as well as aging, thus characterized by an expanding group of care recipients with an intensifying care need per capita. In addition, a significant proportion of out-of-hours contacts are deemed medically non-urgent, constituting a tendency of unnecessary use and overdemand (1–4).

Such superfluous utilization of the available care systems could lead to an unsustainable workload. These factors further burden healthcare providers, increasing the risk of poor patient outcomes (5).

Furthermore, COVID-19 has shown that a sudden increase in healthcare seeking behavior quickly overloads a system that under normal circumstances already operates close to its maximum capacity. Additionally, a pandemic forces a minimization of face-to-face contacts adding to the need for additional pathways into the care system (1).

Overconsumption of OOH-care also comes at an increased cost, further straining the affordability of the healthcare system. Cost-effective interventions to safeguard its sustainability are therefore of primordial importance (3).

Potential demand management strategies were tested to mitigate these challenges. Co-payment, online advice, an overview of the medical cost, and a GP appointment the next morning were all investigated as measures to influence parents' decision process in OOH-care for their children. Online advice was reported to be the only intervention that could potentially affect healthcare seeking behavior in both medically urgent and non-urgent cases without limiting the accessibility of care (6).

Online medical advice fits within a larger trend as searching the internet for health information and potential diagnoses is an increasingly common phenomenon. In Australia about 80% of the population uses the internet for health concerns and 40% searches for self-treatment advice (7). Similar results were found in the US, where about 33% of residents reports attempting to self-diagnose their symptoms via online research (8). This often serves as a precursor to a medical consultation in an attempt to assess severity and thus urgency, with the majority of UK adults consulting the internet beforehand (9).

1.1. Objective

Currently there are a multitude of online symptom checkers and self-triage tools available to the public. These are driven by different entities, such as government bodies with NHS 111 Online in the UK and SNS24 in Portugal, professional associations or commercial firms. However, they are often found to lack a solid evidence-based foundation concomitantly undergoing varying degrees of often self-developed validation. This scoping review aims to provide an overview of the current state of the art of this rapidly progressing field, covering user demographics, safety, accuracy, compliance, cost effectiveness, impact, user experience and complementary implementation with other demand management measures. In doing so, it endeavors to map existing limitations and gaps in the supporting evidence, providing a guide for researchers to conduct relevant studies.

2. Methods

The study design of our scoping review is based on Arksey and O'Malley's (10) five-stage methodological framework.

2.1. Sources and search strategy

A comprehensive search strategy was developed to meet the stated objectives. It is based on three defining features that are reflected in the search terms, highlighting the necessity of having a triage and/or diagnostic function, being exclusively patient-operated, and available in a digital format. The final search strategies can be found in **Supplementary Appendix 1**.

It was subsequently applied to six databases: Cochrane library, NICE, DARE, NIHR, Pubmed, and Web of Science.

2.2. Selection criteria

The search strategy was not filtered on publication date or language, thus encompassing all published works up until July 15, 2022. Studies had to be conducted within a developed health care setting.

Guidelines, primary studies reporting an outcome, and literature reviews were considered for inclusion. Conference abstracts, presentations, opinion pieces, editorials, and comments were excluded. Studies suffering from conflicts of interest were also discarded.

This scoping review aims to cover current evidence on patient-facing digital tools that are text-based and cover a broad range of complaints and medical conditions. Thus, algorithms were excluded if they catered to health professionals, only applied to specific conditions (e.g., COVID-19), focussed on specific situations (e.g., disaster triage). Systems that required additional intervention, such as pictures, video-calls, and any other variation of teleconsultation were also excluded.

2.3. Study selection

Articles were screened on title and abstract after elimination of duplicates, keeping the defining characteristics in mind. Subsequently, two researchers independently performed full text evaluations of the articles considered eligible in the first stage. Periodic discussions were held throughout the review process to reach conformity in case of discordance. Additionally, the references of selected articles were screened.

2.4. Data charting and data extraction

Two authors performed data extraction and charting in duplicate in accordance with a predetermined protocol. The following variables were extracted: author, country, publication date, type of study, methodology, sample size, outcomes assessed, and major findings.

3. Results

A total of 10,250 articles were identified through the initial search strategy. Screening of title and abstract excluded 10,207 articles. Full text review was carried out for 43 articles. 28 studies conformed to the eligibility criteria and were included in the literature review. This process is illustrated in **Figure 1**.

The retained studies originated from the UK (6), the USA (6), Canada (3), the Netherlands (3), Australia (2), (Germany (2), Finland (1), Hong Kong SAR, China (1), New-Zealand (1) Norway (1), Russia (1), and Thailand (1).

Primary research comprised five cross-sectional surveys, five mixed method studies, five audit studies, one prospective cohort study, one retrospective cohort study, one retrospective observational study, one survey study, one population-based descriptive study, one user study, one critical analysis, and one retrospective analysis. The literature reviews that were withheld consisted of two systematic reviews, two scoping reviews, and one literature review. **Table 1** provides a chronological overview of the included studies and the different topics they cover.

3.1. Study outcomes

3.1.1. Demographics

Online symptom checkers and similar services generally have a diverse user base. However, participants are more likely to be younger, female (3, 11) and more highly educated (12). Other factors, such as having low health literacy (irrespective of education level), high technology literacy (13), or limited access to care (14) also seem to increase the likelihood of using a symptom checker. Interestingly, evidence suggests that having a condition that is perceived as stigmatizing, awkward or sensitive contributes to use of these tools (3, 13, 14), indicating that such systems could potentially lower the threshold to seek care for more intimate medical problems.

3.1.2. Safety

Digital triage services are generally considered risk averse (15–17), favoring sensitivity over specificity, often over-triaging, assigning a higher urgency level than is required. A systematic review by Chambers et al. (12) concluded that the available studies reported no evidence of a detrimental effect on patient safety in both simulated and real settings. Important to note is that the strength of evidence is considered weak and thus insufficient, in part due to the limited number of adverse effects reported (18).

More recent studies note a general evolution toward less risk averse triage behavior in symptom checkers, in an effort to offer more suitable advice in low-acuity cases. However, this currently seems to impact safety, with a decreased sensitivity toward more urgent conditions and consequently more missed emergencies (19–21). One study building in part on the foundational research by Semigran et al. (16) finds an average of $>40\%$ of emergencies undertriaged by a collection of 22 systems (21).

3.1.3. Accuracy

The evaluation of accuracy should be divided into two categories. Diagnostic accuracy pertains to digital services that provide a list of potential diagnoses ranked by likelihood and conformity to the clinical picture. Triage accuracy gages the precision of assigned urgency levels by these tools.

Triage accuracy may be the more significant metric, supporting the notion that triage should be the primary function of these tools. Making sure people seek appropriate care can contribute more to their health than attempting to identify the specific origin of their care need (16, 22, 23). A sentiment corroborated by patients, believing self-triage to be more useful to them (13).

3.1.3.1. Diagnostic accuracy

A landmark study by Semigran et al. (16) in 2015 examining 23 symptom checkers found overall diagnostic accuracy to be lacking, with a correct primary diagnosis in only 34% of cases and within the top 20 suggested diagnoses in 58% of cases. Subsequent studies and literature reviews echoed the generally

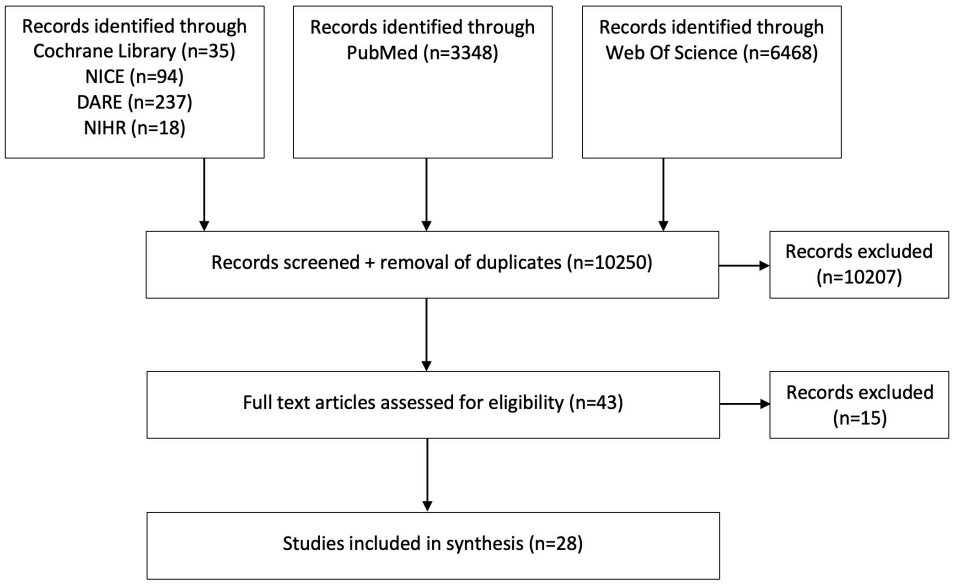


FIGURE 1
Study selection flow diagram.

poor diagnostic performance of these algorithms and the sparse evidence surrounding it (12, 15, 20).

Furthermore, accuracy varied widely depending on the platform, setting, user and disease tested (14, 15). Performance is notably better for common conditions than rare diagnoses. Moreover, women and more highly educated users appear to be more successful at selecting their condition out of the line-up of probable diagnoses (14).

3.1.3.2. Triage accuracy

Results on the accuracy of digital triage tools were more mixed. Yet, on average, these algorithms performed sub-optimally, owing in part to their risk averse nature (20, 24, 25). Even throughout the past years there generally has not been a markable improvement in signposting accuracy of these algorithms. More so, there is evidence that digital triage tool performance increasingly likens the triage decisions of a layperson, including their mistakes (21).

Important to note, however, is the significant variability in triage accuracy between the different tools. A comparative study reported values ranging from 33% to 78%, yet did not manage to reveal a decisively valid tool (16). Even though the field generally appears to be inadequate and seems to lack the capacity to evolve, some symptom checkers offer more promising results, portraying superior sensitivity, specificity, and accuracy (3, 19–21, 24, 26), even advancing in their capabilities over time (20). Such tools have the potential to effectively contribute to patients' healthcare seeking behavior in some cases (20, 21, 24).

Certain characteristics of triage systems were found to be beneficiary to triage accuracy. A review of 36 symptom checkers and triage tools in Australia by Hill et al. (22) concluded that triage tools that take into account demographic data were more accurate than their counterparts. The use of AI algorithms also appeared to benefit performance. Furthermore Verzantvoort et al. (3) reported that tools developed by physician organizations showed above average results, as opposed to those created by a commercial entity or a government.

When examining these results, it remains primordial to consider the limited comparability across studies and systems due to significant heterogeneity of study designs, interventions, and measured endpoints (1, 12, 27). Moreover, the majority of algorithms was examined using clinical vignettes, limiting the validity as well as transferability to the real-world setting (15, 16, 21, 24).

3.1.4. Compliance

There is very limited evidence available on compliance of patients that utilize digital triage tools (12, 16). Overall, patients seemed relatively inclined to follow the proposed advice, with studies finding 57–67,5% of participants to be compliant (3, 11, 28). Although users of NHS111 Online were less likely to comply than those that utilized telephone triage (67.5% vs. 88%; $p < 0.001$).

It has been reported that people were more motivated to initially seek primary care or self-management instructions after being advised to contact the emergency number or visit the ED (28), often because these recommendations were perceived as inappropriate and unnecessary owing to the risk averse nature of the algorithms (11).

10.3389/fmed.2022.1040926

TABLE 1 Included studies of online symptom checkers and triage tools ordered chronologically.

Article information					Topics								
References	Date	Country	Type	Sample size (*n*)	User demo-graphic	Safety	Diagnostic accuracy	Triage accuracy	Compliance	Impact	User experience	Cost effectiveness	Combination of demand management strategies
Schmieding et al. (21)	May 2022	Germany	Follow-up audit study	22 symptom checkers, 45 case vignettes		X		X					
Kujala et al. (36)	May 2022	Finland	Cross-sectional survey	639 health care professionals							X		
Tsai et al. (35)	March 2022	USA	Mixed method study: interview study and user study	Interviews *n* = 25 Users *n* = 20							X		
Dickson et al. (17)	February 2022	UK	Retrospective cohort study	25.333 self-assessments				X					
Chan et al. (19)	December 2021	Canada	Prospective cohort study	281 hospital patients, 300 clinic patients		X		X		X			
Turner et al. (11)	November 2021	UK	Mixed method study	Qualitative survey user questionnaire: Telephone *n* = 795. Online *n* = 3.728. Interview *n* = 32. Staff interview *n* = 16.	X				X	X	X	X	X
Schmieding et al. (21)	July 2021	UK	Observational audit study	12 symptom checkers, 50 case vignettes		X	X	X		X			
Yu et al. (25)	March 2021	Germany	Survey study	15 symptom checkers, 45 case vignettes, 91 users				X					

(Continued)

10.3389/fmed.2022.1040926

TABLE 1 (Continued)

Article information					Topics								
References	Date	Country	Type	Sample size (*n*)	User demo-graphic	Safety	Diagnostic accuracy	Triage accuracy	Compliance	Impact	User experience	Cost effectiveness	Combination of demand management strategies
You et al. (34)	January 2021	USA	Mixed method study	10 semi-structured interviews, 2000 user reviews							X		
Aboueid et al. (13)	January 2021	Canada	Qualitative cross-sectional survey	Survey *n* = 1.547	X						X		
Cross et al. (23)	January 2021	Australia	User study	512 simulated self-assessments							X		
Morse et al. (1)	November 2020	USA	Population-based descriptive study	26.646 self-assessments	X			X					
Schmieding et al. (24)	July 2020	Hong Kong SAR, China	Audit study	2 symptom checkers, 100 ED patient records				X					X
Hill et al. (22)	June 2020	Australia	Follow-up audit study	1.170 diagnosis vignette tests, 688 triage vignette tests			X	X		X			
Gottliebsen et al. (15)	May 2020	Norway and Sweden	Literature review	17 publications		X	X	X		X			
Sutham et al. (26)	April 2020	Thailand	Mixed method study	12 emergency physicians				X			X		X
Donovan et al. (31)	February 2020	UK	Systematic review	3 publications		X		X		X	X		
Meyer et al. (27)	January 2020	USA	Cross-sectional survey	Survey *n* = 329	X			X		X	X		
Chambers et al. (12)	August 2019	UK	Systematic review	29 publications	X	X	X	X	X	X	X	X	

(Continued)

10.3389/fmed.2022.1040926

TABLE 1 (Continued)

Article information					Topics								
References	Date	Country	Type	Sample size (*n*)	User demo-graphic	Safety	Diagnostic accuracy	Triage accuracy	Compliance	Impact	User experience	Cost effectiveness	Combination of demand management strategies
Aboueid et al. (14)	May 2019	Canada	Scoping review	19 publications	X		X						
Millenson et al. (18)	September 2018	USA	Scoping review	30 publications		X							
Verzantvoort et al. (3)	June 2018	Netherlands	A prospective, cross-sectional study	Questionnaire: online *n* = 4.456 Phone *n* = 126	X		X	X	X	X	X		
Polynskaya et al. (37)	June 2018	Russia	Mixed method study	Survey *n* = 200 Interview *n* = 40 Focus group = 1							X		
Giesen et al. (6)	July 2017	Netherlands	Cross-sectional survey	Survey *n* = 377, 1.367 cases						X			
Elliot et al. (32)	December 2015	UK	Retrospective observational study	3.37 million self-assessments						X			
Semigran et al. (16)	May 2015	USA	Audit study	23 symptom checkers, 45 case vignettes		X	X	X					
Lupton et al. (29)	May 2015	New Zealand	Critical analysis	35 symptom checkers						X			
Nijland et al. (28)	2010	Netherlands	Retrospective analysis	6.538 self-assessments Survey *n* = 192	X				X				

Topics marked with "X" are covered by the corresponding study.

In a study by Verzantvoort et al. (3) 65% of patients reported that they intended to comply with the outcome of that specific triage tool. People were more likely to follow the advice when urged to contact their own GP during office hours (75%), followed by self-care advice (67%), OOH-care (61%) and wait-and-see instructions (56%). Certain patient characteristics correlated with a higher compliance, such as <13 years (OR 1.8, 95% CI: 1.3–2.3, $p < 0.001$), male sex (OR 1.2, 95% CI: 1.1–1.4, $p = 0.045$), and user satisfaction (OR 2.5, 95% CI: 2.2–2.9, $p < 0.001$). Main reasons for defying the advice given related to inability to accurately convey complaints, contact with a physician pre-dating triage and preferring their own judgment. An important limitation of this study lies in the fact that it only quantifies intention to comply, instead of the resulting care seeking behavior (3).

3.1.5. Impact

3.1.5.1. Impact on health literacy

Availability of online symptom checkers and triage tools can be a valid source of health information, allowing patients to educate themselves (29) and gain further insights in their health status and conditions (27). An increase in health literacy could subsequently benefit the patient-physician relationship (22).

3.1.5.2. Impact on health equity and digital divide

As established, demographics of users of symptom checkers and triage tools tend to show a younger, more educated user base. Older patients, or those less educated utilized telephone triage and direct contact more often. This could potentially affect health equity (12). In contrast, a study by Morse et al. (1) on use characteristics of a digital symptom checker found a significant proportion of patients to be of older age, belonging to a subpopulation not typically associated with regular use of online resources. These mixed results underline the need to evaluate a tool in its intended setting and population.

3.1.5.3. Impact on OOH-care

Available research on the impact of triage systems on OOH-care is scarce.

3.1.5.3.1. Healthcare seeking behavior

It has been reported that digital triage tools and symptom checkers have a limited ability to modify health care seeking behavior by informing patients and assisting them to make medically appropriate decisions. It was found that such guidance could potentially improve the safety of parent's decisions in the management of children with possibly severe ailments (6). However, research was often focussed on specific conditions or settings and offered indirect evidence, limiting the ecological validity of their conclusions (12).

3.1.5.3.2. Workload

Furthermore, evaluation of the impact of triage algorithms on the burden of the health care systems shows highly variable results. Some studies report a potential decrease in pressure on the health care system (3, 30). Which could, by extension, reduce the urgency of current health care staffing shortages (1).

Others, however, did not observe an effect on workload, as Donovan et al. (31) were unable to discern a digital intervention capable of altering urgent care usage, based on available data.

More importantly, the majority of studies that reported on this topic has voiced concerns that symptom checkers and triage tools might conversely increase inappropriate OOH-care use due to their risk-averse nature, often advising additional health care interaction, even though self-care or a wait and see policy would be adequate (11, 15, 20).

3.1.5.4. Accessibility to care

It is relevant to note, however, that a growing demand and rise in health care utilization does not necessarily equate to an increase in inappropriate use. Implementation of NHS 111 online resulted in a significant new demand, with people finding their way to the available health care providers more easily. This suggests that symptom checkers could increase accessibility to care, lowering the threshold for those in need (11, 22). More so, people were found to be more motivated to seek medical care when assisted in their decision by a tool (12).

The inconsistent picture these studies paint of the impact of available systems on OOH-care, highlights the importance of studying a tool in its intended real-life setting to accurately assess its influence. Additionally, it will be necessary to monitor a potential shift to more appropriate use, with people previously unaware of their need of care finding their way to OOH-care services more easily, thereby likely clouding certain established outcome parameters such as workload (11, 22).

3.1.5.5. Impact on public health

The data collected by such tools could also impact public health management by providing direct epidemiological data that can be used to map the evolution of infectious diseases (32). Online data generated by symptom checkers was found to capture evolutions earlier than traditional surveillance or telephone triage could. Additionally, it is able to provide insights on symptomatic patients that do not contact the health care system directly. Thus, diagnostic and triage algorithms have the potential to serve as a complementary source within national surveillance systems, especially during crises such as the COVID-19 pandemic (33).

3.1.6. Cost effectiveness

There is very limited evidence available on the effect of symptom checkers and triage tools on costs. In an evaluation of NHS 111 Online during the implementation phase, costs were lower compared to NHS 111 telephone triage, potentially in part due to the on average lower acuity of complaints processed. When both systems operated simultaneously, a shift of ≥38% of telephone contacts to digital triage would be necessary to achieve a cost reduction (11).

Two small studies reported significant cost efficiency both in operational expenses as well as in care diversion. However, these savings were self-reported and considered inadequate to come to a consensus (12).

3.1.7. User experience

3.1.7.1. Patients

A systematic review by Chambers et al. in 2019 retained 9 studies investigating patient and/or caregiver satisfaction. Patients predominantly considered the examined symptom checkers and triage systems to be very satisfactory (12). This sentiment was corroborated by Meyer et al. (27) in their analysis of patient perspectives on the usefulness of these services. They reported patients finding the tool easy to use and useful, and made mention of a general willingness to use the tool on a recurring basis (27).

More recent research nuances this general trend to a degree. NHS 111 users were found to be more satisfied with the telephone service than the online equivalent (50% vs. 71%; $p < 0.001$) (11). Other studies showed that some prefer traditional search engines over set algorithms, mostly due to the perception that they offer more freedom in describing their symptoms and providing information (13, 23).

Several factors were observed as beneficial to usage, such as a limited accessibility to care, backing of these tools by credible sources, such as government entities and caregiver associations, or integration in the care system. Conversely, there are still multiple obstacles such as restricted internet access, the use of medical jargon, as well as reservations about data privacy and trust (13).

To overcome these hurdles and improve user experiences, several advancements have been suggested. On the one hand, a more customizable input would increase the perceived flexibility of the system, allowing patients to describe their pattern of symptoms more accurately and thus feel more heard. On the other hand, tools should be made more comprehensible and allow the patient insight into the decision-making process (34). This can be achieved by using unambiguous language and offering comprehensive explanations (13, 34, 35).

Interestingly, one study observed a significant lack of awareness of symptom checkers, with >50% of participants unfamiliar with the technology, which severely limits usage. To optimize the use of these tools, targeted interventions could be implemented, tailored to individual subpopulations of potential users (13). Furthermore, support and recommendation by health care providers and credible associations has the potential to impact use considerably (13, 35, 36).

3.1.7.2. Caregivers

Evidence on the experiences of healthcare providers with symptom checkers is sparser. Current studies report a generally positive perception of the algorithms, with health care providers believing it could be beneficial to both patients, as well as caregivers (36).

Anticipated benefits relate to a potential reduction in workload, while simultaneously offering a more expeditious, accessible and supportive service to patients. Triage services were believed to be especially useful during the pandemic, when demand increased significantly (36).

Several challenges were acknowledged pertaining to the suboptimal impact on workflow with risk of multiple channels of contact per patient, the inaccuracy of triage, the effect on the digital divide, and a perceived threat to professional autonomy (36). Even so, digital diagnostics are considered a part of the future of medicine by many caregivers (37).

User experience is naturally specific to any individual tool. The overall trend in the above-mentioned reports, however, is more consistent and expected to be more generalizable (12).

3.1.8. Combination of demand management measures

The implementation of an online self-triage service is one possible measure to counter the current challenges facing OOH-care systems. There is some evidence on the effects of combining these tools with other interventions, such as telephone triage.

Multiple studies reported that performance of algorithms was comparable to that of existing telephone triage services, such as NHS111 (19, 32). Parallel operation of both systems did not lead to significant reductions in telephone contacts during the initial phase, as some interacted with both, to confirm their findings (11). There are, however, indications that implementation of triage services as a complementary service to telephone triage could be beneficial to the healthcare system, as it could lower the threshold for care and offer a suitable alternative approach for lower acuity and non-trauma problems (11, 25, 26); a potential advantage echoed by telephone staff of NHS111 (11).

4. Discussion

A 2019 systematic review by Chambers et al. concluded that the research available at that time was considered weak, the majority consisting of observational studies, and clouded by an abundance of gray literature (12). In recent years, there has been a notable increase in research surrounding the technology as well as ehealth in general. However, there are still several limitations, both pertaining to the available tools, as well as the studies evaluating them.

Research examining safety and accuracy of these systems highlights the persistence of a risk-averse disposition (12, 15–17). An effort is being made to evolve to more balanced algorithms, but some studies report that this evolution currently comes at the expense of the technology's safety, which should always remain their priority (19–21). More so, these developments seem to miss their mark at present, with most

systems still appearing to be insufficiently accurate and show limited progress to date.

The applied methodologies to evaluate these parameters are suboptimal as well, with most of the available studies relying solely on clinical vignettes. Such simulations are a good technique to benchmark and compare different tools, offering an initial insight in their performance. They are, however, insufficient to determine functionality in a real-life setting and should serve exclusively as a basis to be supplemented by studies in the intended environment. Current research fitting these requirements was often limited by a relatively small study sample and short duration, or restricted to certain conditions and subpopulations. A finding that echoes the results of previous studies (12, 15, 34).

Future research should be implemented in the real world and be of a larger scale and scope, allowing for continuous, multifaceted data collection to monitor foundational aspects, such as safety and accuracy, and compare them against the gold standard of a medical evaluation. This would require access to personal medical information, something research groups often lacked (1, 27).

The same applies to other case-specific outcomes such as cost-effectiveness, a topic that currently remains unclear, with a handful of self-reported and non-peer reviewed studies being regarded as inadequate. An initial evaluation of NHS 111 Online did show a potential economic benefit if enough patients shifted from telephone to online triage (11). This evidence is not transferable due to the significant multifactorial influences making the results highly specific to the circumstances.

We concur with the statement that triage should be the most prominent area in the future of these systems, as it has the most potential to have a significant impact on both individual patients, and the health care system (16, 23).

On a positive note, there is more consensus regarding the experiences of users and health care professionals. Patients generally found the tool user-friendly, usable, and useful (12, 27).

Healthcare professionals too believe in their potential benefit to both patients and caregivers (36) and reportedly envision a role for it in the future of medicine (37). There are, however, still some hurdles to overcome, such as the use of medical jargon and questions surrounding privacy, which hinder trust (13). These findings are expected to be more generalizable to different tools and settings (12).

To further optimize patient experiences, tools should become increasingly more patient-centered. This can be achieved through a transparent policy, understandable language and comprehensive explanations (13, 34, 35), thus further empowering patients and enhancing trust. Ideally, patients should be included in the design process of these tools.

Evidence on compliance was scarcer, finding users to be relatively inclined to follow the guidance given (3, 11, 28), but often apprehensive to follow through because of advice perceived as excessive owing to the risk-averse reactions of most tools (11, 28). The above-mentioned research mostly measured intention to comply, rather than real behavior, severely limiting the strength of evidence. Compliance should therefore be investigated by tracking patient flow and measuring healthcare seeking habits in a real life setting. In addition, the impact of factors influencing compliance, should be mapped.

The impact of tools on the workload remains uncertain, with highly variable results being reported. The top-performing apps have the potential to influence healthcare seeking behavior toward appropriate care, with both patients and the healthcare system benefitting (3, 6, 30). However, multiple studies demonstrated how the risk-averse disposition of some tools could conversely lead to an increase in healthcare demand, thus foregoing its value in this regard (11, 15, 20).

Evidence suggests, however, that their role reaches further than demand management. Digital triage tools could open a new avenue to connect with the healthcare system, lowering the threshold to some and improving the accessibility of care (11, 22). Moreover, with patients increasingly turning toward the internet for health information, there is a clear need for validated information, which symptom checkers could provide through tailored advice, potentially positively impacting health literacy in the process (6, 12). Conversely, concerns have been voiced that preliminary self-diagnosis through these systems could contribute to increased anxiety in some patients (16). These conflicting statements warrant additional research on patient perspectives.

There is some concern about the impact of this technology on the existing digital divide and the potential consequences for health equity. Current research reports a user base skewed more toward those younger, more educated and technologically literate (3, 11–13), potentially posing a challenge for certain subpopulations, such as the elderly. Implementation of these tools should therefore be as an adjunct to other channels, rather than as a replacement (11, 25, 26).

Additionally, a potential advantage of these tools in crises was illustrated by their use in the pandemic. A veritable plethora of tools was created within a short time frame to help curb the exponential demand in healthcare seeking, as became apparent during our search strategy. They served a double function, unburdening the healthcare system by providing trustworthy background information and guidance in times of an information overflow (36), as well as collecting epidemiological data for national surveillance (33).

We therefore recommend that clear objectives be set prior to implementing and researching the tools, allowing for relevant and appropriate outcomes to be studied. This approach facilitates a more comprehensive understanding of the impact on the healthcare system, including the effects on demand management, accessibility of care and patient education. It is this all-encompassing analysis that can determine the ultimate added value of the technology.

The primary challenge for the field remains the significant heterogeneity across studies and triage systems. The tools were shown to vary significantly in terms of functionality, performance, and triage approach. The additional variety in study designs, interventions, quality and measured endpoints of current literature further limits the generalizability of results. This observation is a confirmation of previous findings, underlining the persistence of this shortcoming (12, 14, 18–20). Algorithms will need to be studied in their intended environment to draw more definitive conclusions.

There are several risks to the field as well, with a multitude of tools flooding a largely unregulated market, often lacking an adequate evidence-based approach and suffering from conflicts of interest. This could be amended by providing independent and transparent research to assist patients and caregivers in identifying the top performing systems (35). Implementation should be conducted within a framework of standardized evaluations to objectively validate triage systems, contributing to regulation of the field (15, 17, 18, 20). Responsible authorities should subsequently provide oversight, recommend validated tools and integrate them into the health care system to optimize functionality and user experience (13, 36).

4.1. Limitations of this review

There are several limitations to this review. It is possible that some potentially eligible articles were not captured by our search method. This could be due to the selection of databases, the applied exclusion criteria or the selected search terms which highlighted a lack of applicable MeSH terms. Furthermore, the absence of a formal quality assessment of the retained studies can make it challenging to accurately appraise the value of reported results to the field.

5. Conclusion

Numerous digital symptom checkers and triage tools are presently available to the public and fit within the trend of an increasing reliance on the internet for access to health information.

The evidence collected in this literature study is characterized by multiple limitations. Nevertheless, with some reservation, several trends can be distilled. Current research highlights the risk-averse nature of these services, which causes challenges for their accuracy. Recent evolutions in algorithms have varying degrees of success.

User satisfaction is generally high, and patients appear to be amenable to the advice given by a digital service, with most participants intending to comply. There is evidence of a multifaceted impact on the healthcare system, with preliminary research seeing potential benefits for accessibility of care, health literacy and syndromic surveillance. In contrast, there is ambiguity about the effects on workload and digital divide, warranting caution.

Notwithstanding these themes, there is a clear need for additional research, with a strong preference for study designs that most closely match the circumstances of the intended definitive setting. Additionally, an evolution toward more homogeneous methodologies, aided by regulation and standardization of evaluations, should increase the generalizability of results, furthering the field as a whole.

Author contributions

AP, HP, and VV designed the review collectively and were responsible for data interpretation and reporting. AP and HP were involved in planning. AP and VV conducted the review. AP attested that all listed authors met authorship criteria and that no others meeting the criteria have been omitted. All authors contributed to the article and approved the submitted version.

Funding

The authors received a grant (number: T000718N) from Fonds Wetenschappelijk Onderzoek (https://www.fwo.be/) for this project, covering the working costs and payment was made to their institutions.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fmed.2022.1040926/full#supplementary-material

References

1. Morse K, Ostberg N, Jones V, Chan A. Use characteristics and triage acuity of a digital symptom checker in a large integrated health system: population-based descriptive study. *J Med Internet Res.* (2020) 22:e20549. doi: 10.2196/20549

2. Gill P, Goldacre M, Mant D, Heneghan C, Thomson A, Seagroatt V, et al. Increase in emergency admissions to hospital for children aged under 15 in England, 1999-2010: national database analysis. *Arch Dis Child.* (2013) 98:328–34. doi: 10.1136/archdischild-2012-302383

3. Verzantvoort N, Teunis T, Verheij T, van der Velden A. Self-triage for acute primary care via a smartphone application: practical, safe and efficient?. *PLoS One.* (2018) 13:e0199284. doi: 10.1371/journal.pone.0199284

4. de Bont E, Lepot J, Hendrix D, Loonen N, Guldemond-Hecker Y, Dinant G, et al. Workload and management of childhood fever at general practice out-of-hours care: an observational cohort study. *BMJ Open.* (2015) 5:e007365. doi: 10.1136/bmjopen-2014-007365

5. Bond K, Ospina M, Blitz S, Afilalo M, Campbell S, Bullard M, et al. Frequency, determinants and impact of overcrowding. *Healthc Q.* (2007) 10:32–40.

6. Giesen M, Keizer E, van de Pol J, Knoben J, Wensing M, Giesen P. The impact of demand management strategies on parents' decision-making for out-of-hours primary care: findings from a survey in The Netherlands. *BMJ Open.* (2017) 7:e014605.

7. Cheng C, Dunn M. Health literacy and the internet: a study on the readability of Australian online health information. *Aust N Z J Public Health.* (2015) 39:309–14. doi: 10.1111/1753-6405.12341

8. North F, Ward W, Varkey P, Tulledge-Scheitel S. Should you search the Internet for information about your acute symptom?. *Telemed J E Health.* (2012) 18:213–8. doi: 10.1089/tmj.2011.0127

9. UK Government ONS. *Internet access in Great Britain, including how many people have internet access, what they use it for and online shopping.* Great Britain, UK: Office for National Statistics (2020).

10. Arksey H, O'Malley L. Scoping studies: towards a methodological framework. *J Soc Res Methodol.* (2005) 8:19–32. doi: 10.1080/1364557032000119616

11. Turner J, Knowles E, Simpson R, Sampson F, Dixon S, Long J, et al. *Health services and delivery research. impact of NHS 111 online on the NHS 111 telephone service and urgent care system: a mixed-methods study*. Southampton, UK: NIHR Journals Library (2021).

12. Chambers D, Cantrell A, Johnson M, Preston L, Baxter S, Booth A, et al. Digital and online symptom checkers and health assessment/triage services for urgent health problems: systematic review. *BMJ Open.* (2019) 9:e027743. doi: 10.1136/bmjopen-2018-027743

13. Aboueid S, Meyer S, Wallace J, 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. doi: 10.2196/22637

14. Aboueid S, Liu R, Desta B, Chaurasia A, Ebrahim S. The use of artificially intelligent self-diagnosing digital platforms by the general public: scoping review. *JMIR Med Inform.* (2019) 7:e13445. doi: 10.2196/13445

15. Gottliebsen K, Petersson G. Limited evidence of benefits of patient operated intelligent primary care triage tools: findings of a literature review. *BMJ Health Care Inform.* (2020) 27:e100114. doi: 10.1136/bmjhci-2019-100114

16. Semigran H, Linder J, Gidengil C, Mehrotra A. Evaluation of symptom checkers for self diagnosis and triage: audit study. *BMJ.* (2015) 351:h3480. doi: 10.1136/bmj.h3480

17. Dickson S, Dewar C, Richardson A, Hunter A, Searle S, Hodgson L. Agreement and validity of electronic patient self-triage (eTriage) with nurse triage in two UK emergency departments: a retrospective study. *Eur J Emerg Med.* (2022) 29:49–55. doi: 10.1097/MEJ.0000000000000863

18. Millenson M, Baldwin J, Zipperer L, Singh H. Beyond Dr. Google: the evidence on consumer-facing digital tools for diagnosis. *Diagnosis.* (2018) 5:95–105. doi: 10.1515/dx-2018-0009

19. Chan F, Lai S, Pieterman M, Richardson L, Singh A, Peters J, et al. Performance of a new symptom checker in patient triage: Canadian cohort study. *PLoS One.* (2021) 16:e0260696. doi: 10.1371/journal.pone.0260696

20. Ceney A, Tolond S, Glowinski A, Marks B, Swift S, Palser T. Accuracy of online symptom checkers and the potential impact on service utilisation. *PLoS One.* (2021) 16:e0254088. doi: 10.1371/journal.pone.0254088

21. Schmieding M, Kopka M, Schmidt K, Schulz-Niethammer S, Balzer F, Feufel M. Triage accuracy of symptom checker apps: 5-year follow-up evaluation. *J Med Internet Res.* (2022) 24:e31810. doi: 10.2196/31810

22. Hill M, Sim M, Mills B. The quality of diagnosis and triage advice provided by free online symptom checkers and apps in Australia. *Med J Aust.* (2020) 212:514–9.

23. Cross S, Mourad A, Zuccon G, Koopman B editors. . Search engines vs. symptom checkers: a comparison of their effectiveness for online health advice. *Proceedings of the web conference 2021*. New York, NY: Association for Computing Machinery (2021). p. 206–16.

24. Schmieding M, Mörgeli R, Schmieding M, Feufel M, Balzer F. Benchmarking triage capability of symptom checkers against that of medical laypersons: survey study. *J Med Internet Res.* (2021) 23:e24475.

25. Yu S, Ma A, Tsang V, Chung L, Leung S, Leung L. Triage accuracy of online symptom checkers for accident and emergency department patients. *Hong Kong J Emerg Med.* (2020) 27:217–22.

26. Sutham K, Khuwuthyakorn P, Thinnukool O. Thailand medical mobile application for patients triage base on criteria based dispatch protocol. *BMC Med Inform Decis Mak.* (2020) 20:66. doi: 10.1186/s12911-020-1075-6

27. Meyer A, Giardina T, Spitzmueller C, Shahid U, Scott T, 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. doi: 10.2196/14679

28. Nijland N, Cranen K, Boer H, van Gemert-Pijnen J, Seydel E. Patient use and compliance with medical advice delivered by a web-based triage system in primary care. *J Telemed Telecare.* (2010) 16:8–11. doi: 10.1258/jtt.2009.001004

29. Lupton D, Jutel A. 'It's like having a physician in your pocket!' A critical analysis of self-diagnosis smartphone apps. *Soc Sci Med.* (2015) 133:128–35. doi: 10.1016/j.socscimed.2015.04.004

30. Poote A, French D, Dale J, Powell J. A study of automated self-assessment in a primary care student health centre setting. *J Telemed Telecare.* (2014) 20:123–7. doi: 10.1177/1357633X14529246

31. Donovan E, Wilcox C, Patel S, Hay A, Little P, Willcox M. Digital interventions for parents of acutely ill children and their treatment-seeking behaviour: a systematic review. *Br J Gen Pract.* (2020) 70:e172–8. doi: 10.3399/bjgp20X708209

32. Elliot A, Kara E, Loveridge P, Bawa Z, Morbey R, Moth M, et al. Internet-based remote health self-checker symptom data as an adjuvant to a national syndromic surveillance system. *Epidemiol Infect.* (2015) 143:3416–22. doi: 10.1017/S0950268815000503

33. Tozzi A, Gesualdo F, Urbani E, Sbenaglia A, Ascione R, Procopio N, et al. Digital surveillance through an online decision support tool for COVID-19 over one year of the pandemic in Italy: observational study. *J Med Internet Res.* (2021) 23:e29556. doi: 10.2196/29556

34. You Y, Gui X. Self-Diagnosis through AI-enabled chatbot-based symptom checkers: user experiences and design considerations. *AMIA Annu Symp Proc.* (2020) 2020:1354–63.

35. Tsai C, You Y, Gui X, Kou Y, Carroll J editors. Exploring and promoting diagnostic transparency and explainability in online symptom checkers. *CHI '21: Proceedings of the 2021 CHI conference on human factors in computing systems*. New York, NY: Association for Computing Machinery (2021). p. 1–17.

36. Kujala S, Hörhammer I. Health care professionals' experiences of web-based symptom checkers for triage: cross-sectional survey study. *J Med Internet Res.* (2022) 24:e33505. doi: 10.2196/33505

37. Polynskaya G, Mesropyan M. Detection of patterns and trends in patient behavior while using electronic applications and Internet resources for self-diagnosis. Бизнес-информатика (2018) 1:28–38.

Digital self-triage: An evaluation of patient compliance

1. Introduction

Several online symptom checkers and digital triage tools are currently available to the public. These systems are part of various demand management strategies aiming to relieve the overburdened health care system by assisting patients in their care seeking behavior, stimulating appropriate usage of the available services in the process.

A scoping review [1] was performed to provide a synthesis of current evidence and identify persisting literature gaps. The emergence of these systems was shown to be accompanied by an intensifying research effort on this subject and on the overarching subject of ehealth. However, current evidence remains characterized by multiple limitations, with notable diversity persisting both in study design as well as tools examined. Triage systems were found to vary significantly in scope, algorithm design and triage approach, while the applied study methodologies differed in context, quality, and measured outcomes. This heterogeneity in the available literature significantly limits transferability of findings [2-5], a concern echoed by previous literature reviews[6, 7].

Notwithstanding these concerns, if the current evidence is framed within the context of these persisting limitations, it is possible to identify certain overarching themes.

Evidence on triage accuracy shows a risk-averse predilection[5, 8, 9], stemming from the logical prioritization of safety over efficiency. Evolution towards more accurate and balanced systems remains difficult[5, 7, 10].

Patients generally report a positive user experience, often finding tools both usable and useful[2, 11]. Overall, caregivers concur, seeing potential benefits for both parties if the evolution towards patient-centeredness continues[12].

Current literature on the subject provides evidence that the effects of this technology could go beyond modifying health care seeking behavior. Initial studies report potential improvements to accessibility of care[13, 14] and the availability of validated medical information, which could benefit the individual patient and improve overall health literacy[2, 15]. Moreover, the data generated by such tools could contribute to national disease surveillance systems[16, 17]. Conversely, the effects on workload[5, 18], digital divide[2, 6, 13] and privacy are more ambiguous and are cause for concern, emphasizing the need for additional research on the multifaceted ramifications of such tools on their intended environment.

The effectiveness of these systems is dependent on patients following the advice given. Current evidence on compliance is scarce, however. Initial results generally indicate a receptive attitude towards generated recommendations[13, 19, 20], although they were frequently perceived as excessive, a result of the risk averse nature of many algorithms, and subsequently devalued by the recipients[1, 13, 20].

It is important to note that these studies predominantly examined intention to comply, rather than quantifying the resulting care seeking behavior. It remains unclear to what extent these outcomes correlate.

Consequently, there is still a lacuna in research regarding the measurement of compliance, which could be achieved by studying patients' actions in a real-life setting and comparing the results to their previously stated intentions.

1.1. Objective

This study aims to quantify patients' health care seeking behavior and subsequently correlate their intention to comply, both pre- and post-advice when using the digital self-triage system moetiknaardedokter.be. Furthermore, factors influencing adherence to the advice given will be mapped. This results in the following research questions:

RQ1: How compliant are people with the advice given by the digital self-triage tool?

- RQ1.1: How do patient, urgency, disease, and triage characteristics influence compliance?

RQ2: To what degree do intention to comply and compliance correlate?

2. Methods

1.2. Study design and setting

A mixed method approach was devised to accurately respond to the research questions posed. A prospective longitudinal observational study was set up, supplemented by a questionnaire, to measure health care seeking behavior in an out-of-hours setting. This examination serves as a pilot study for a more comprehensive research project in the context of a doctoral dissertation performed by author A.P.

1.3. The tool

Moetiknaardedokter.be is a digital self-triage tool based on the Nederlandse Triage Standaard (NTS) and NHG (Nederlands Huisartsen Genootschap) guidelines. The service is available in eight languages: Dutch, English, German, Arabic, Turkish, Polish, Ukrainian and Russian. Patients initially provide basic demographic characteristics by specifying sex at birth and date of birth. Subsequently, the affected body area is selected, and a preset list of complaints is provided to help

Table 1. Triage levels of urgency

	Out-Of-Hours advice	Advice during office hours
U1 – Life-threatening	Call 112 or the GPC immediately.	Call 112 or your family doctor immediately.
U2 – Emergency	Call the GPC immediately.	Call your family doctor immediately.
U3 – Urgent	Call the GPC. Contact the dentist on call.	Call your family doctor. Call your dentist.
U4 – Routine	Call the GPC.	Make an appointment with your family doctor. Contact your family doctor.
U5 – Other day / Self-care advice	There is no need to see a doctor. Contact the dentist on call.	There is no need to see a doctor. Contact your dentist.

define the primary ailment. Users are posed a series of yes/no questions screening for potentially life-threatening conditions in decreasing order of severity. Patients are then assigned to one of five urgency categories, receiving the corresponding recommended course of action applicable to that specific time. This recommendation is generally followed by care advice, either in anticipation of a medical consultation or as part of self-treatment. (Table 1)

1.4. Data

The triage algorithm was made available to the public via integration into the website of a General Practitioner Cooperative (GPC). Data of every patient utilizing the triage system in the participating GPC region from 01/03/2023 until 31/03/2023 was examined for inclusion into the study. In accordance with Belgian medical and GDPR legislation, patients were offered the opportunity to exclude their anonymous data from scientific research by ticking the opt-out box. Approval of the ethics committee of UZA-UAntwerp was obtained on 13/12/2021. (Appendix 1) Estimation of a representative sample size was based on a confidence interval of 95%, a margin of error of 5% and potential total population size of 6.768.135, comprising the population of Flanders(2022).

Triage results were limited to those that potentially overlapped with the opening hours of the specified GPC. Thus, triages completed between 8 AM and 4 PM were excluded from the study. In the instances when patients performed multiple sequential triages, the outcome with the highest level of urgency was withheld. If every triage resulted in the same level of urgency, the most recent triage was included. To accurately measure compliance, every outcome recommended by the algorithm should be verifiable through a secondary source. Since U1 recommendations, considered life-threatening, receive the advice to either call 112 or the GPC immediately, they were excluded from the study, as call-data of the corresponding 112 call center was not available for comparison at the time of data collection.

Additionally, patients utilizing the triage tool for dental complaints were excluded as well, as the tool advises patients requiring medical attention to contact the dentist on call, of which there is no centralized dataset.

Routine registration data was collected from the algorithm anonymously, including date, time of triage, age, sex, postal code (if provided voluntarily), affected region of body, complaint category, grade of urgency, triage advice, allocated device ID and session ID, triage duration, triage language, and originating GPC website or postal code. Entries missing any of the variables were excluded from the dataset.

Moreover, a questionnaire was developed to gain insight into patients' intentions both prior to and after receiving the generated advice. The survey also inquired about the motivational drivers behind compliance or non-compliance with the recommendation provided, as well as the tool's ability to modify health care seeking behavior. Patients completing a triage session were offered the option to voluntarily take part in a short questionnaire for scientific purposes, which was located below the generated triage advice. The following questions were posed:

1. Did you think you needed to see a doctor before completing this tool?

- Yes
- No

2. Do you intend to follow the advice given?

- Yes → Continue to Q3a
- No → Continue to Q3b

3a. Why do you/will you follow the advice? Continue to Q4

- I trust the advice.
- I am sufficiently informed.
- I am reassured.
- Other: Free text option

3b. Why not?

- I don't trust the advice.
- I need more information.
- I am not comfortable yet.
- I have not been able to explain my situation.
- Other: Free text option

4. Did the advice you have received change your mind?

- Yes

- No

5. Do you feel like you were able to sufficiently clarify your complaints?

- Yes

- No

Subsequently, the routine registration data from the algorithm was compared to the database of the iCAREdata-project. This central clinical research archive contains records of every patient interaction with the out-of-hours care services of the studied regions, linking interdisciplinary information between different services while adhering to the current requirements on privacy, ethics, and legislation.

iCAREdata provided the following routine registration data: date, time of contact, birth year, sex, postal code, method of contact, type of contact, complaint (ICPC-2), diagnosis (ICPC-2), grade of urgency, possible referral, allocated contact ID, and anonymized social security number.

Data-matching was achieved through a two-step process. A dedicated algorithm was developed to link the datasets on five anonymous variables: age, sex, time of triage and contact up to an interval of 3 hours, complaint category, and region. Doubles were then examined manually, taking in account time of contact, complaint category, and diagnosis. In case of persisting ambiguity, all implicated pairings were excluded.

1.5. Outcomes

The resulting dataset visualizes the patient flow through the health care system after self-triage. At the individual level, it is possible to compare the decisions made and actions taken to the self-triage assessment and corresponding advice, thus measuring actual compliance. Furthermore, influencing and triage characteristics (complaint category, level of urgency, time of day). By also taking into consideration the information gathered by the questionnaire, a comparison can be made between intention to comply and adherence to the advice given, including the effects of influencing variables.

1.6. Statistical analysis

Descriptive statistics were performed to report several variables of the triage data and survey responses. Demographics of triage users, survey respondents, and patients contacting the GPC service out-of-hours were detailed and compared. Subsequently, overall compliance was examined and correlated to the reported intention to comply and potentially induced behavioral change. Reasons listed by survey respondents motivating their intention were described as well.

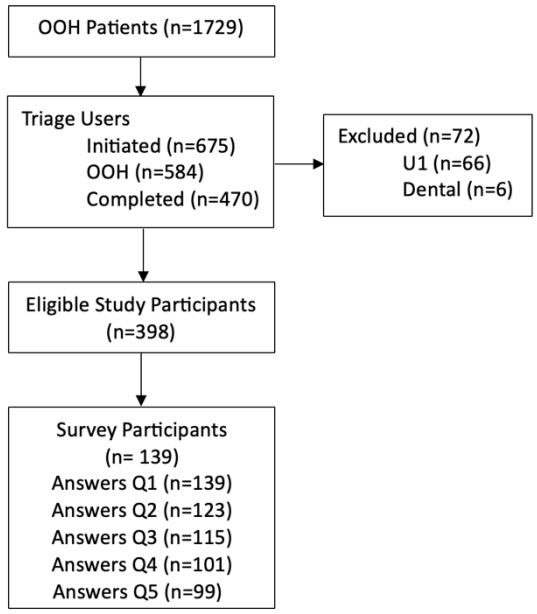
To analyze the factors influencing compliance, multivariate logistic regression analysis was performed. The following variables were examined: age, sex, assigned urgency category, and listed complaint category. IBM SPSS Statistics version 28 was used to for all statistical analysis. A p-value <0.05 was regarded as statistically significant.

3. Results

3.1. Participants

Over the study period 675 people used the self-triage tool provided on the website of the participating GPC, of which 584 occurred during the predetermined time periods outside office hours. Out of these, 470 participants completed a triage protocol, yielding a conversion rate of 80,5%. 66 U1 triage results and 6 dental complaints were excluded in accordance with the predefined criteria. The resulting 398 triage results were included in the study. 139 self-triage users subsequently initiated the accompanying survey, with 99 respondents completing it, resulting in a full response rate of 24,9%. The drop-off rate per question was visualized in the participants flow diagram (Figure 1). Concurrently 1729 patients contacted the participating GPC during the study period.

Figure 1. Flow diagram of participants



3.2. User demographics

The demographic data of the included triage users shows that there is a widespread age distribution. There is a notable peak in triages for patients under 5 years of age (11.6%), which is similar in proportion to the observed number of young patients who were found to contact the GPC (12.1%). In contrast, we see a greater share of younger patients using the digital triage system, with adults between 20 and 40 years of age making up 41.7% of the total userbase, compared to 26.9% of the patients consulting the GPC. In line with these results, the

number of users decreases significantly with age, with those 75+ of age making up only 1.1% of the total triage userbase. By comparison, GPC users in the same age category were found to total 14.3% of contacts. (Table 2) (Figure 2)

Survey respondents were found to be relatively in line with triage users in terms of age distribution, albeit with a relatively higher proportion of people between the ages of 40 and 60 (31.3%), as opposed to triage users (23.4%). (Table 2) (Figure 2) 56.5% of triage users were female, which is in line with the percentage of female GPC users (54,5%) and survey participants (56,5%) for this study. (Table 2)

Table 2. User Characteristics

Variable	% of GPC users (n = 1729)	% of triage users (n = 398)	% of survey participants (n = 99)
Age			
0≤UTAI<5	12.1%	11.6%	9.1%
5≤UTAI<10	5.5%	5.8%	4.0%
10≤UTAI<15	4.8%	4.5%	3.0%
15≤UTAI<20	5.3%	5.8%	6.1%
20≤UTAI<25	6.7%	10.8%	7.1%
25≤UTAI<30	7.0%	13.1%	8.1%
30≤UTAI<35	6.9%	8.8%	10.1%
35≤UTAI<40	6.3%	9.0%	10.1%
40≤UTAI<45	4.5%	5.3%	6.1%
45≤UTAI<50	5.4%	5.3%	10.1%
50≤UTAI<55	4.8%	6.3%	6.1%
55≤UTAI<60	5.3%	6.5%	9.1%
60≤UTAI<65	4.8%	3.0%	4.0%
65≤UTAI<70	3.6%	1.8%	3.0%
70≤UTAI<75	2.7%	1.5%	3.0%
75≤UTAI<80	3.9%	0.5%	1.0%
80≤UTAI<85	3.5%	0.3%	
85≤UTAI<90	3.7%	0.3%	
90≤UTAI<95	2.5%		
95≤UTAI<100	0.7%		
100≤UTAI<105	0.1%		
Sex			
Female	54.5%	56.5%	57.6%
Male	45.5%	43.5%	42.4%

Distribution of age and sex in percentages per examined user group.

3.3. Triage characteristics

Triage users most frequently sought advice regarding coughing and ear, nose, and throat (ENT) complaints. Skin and eye complaints complete a top five dominated by symptoms related to infectious diseases (Table 3).

In 12,1% of included triage results, users were advised not to contact the GPC system and received self-care advice or were recommend wait-and-seeing. 79,1% were considered high(er) acuity cases requiring a timelier medical assessment. The U1-triages, considered life-threatening were excluded from the assessment of compliance, but made up 66 of 464 originally completed protocols, accounting for a provisional 14.2% of generated advice.

A significant majority of triages were conducted in the weekend (79.4%), as opposed to during weeknights (20.6%).

Table 3. Triage characteristics

Top 5 complaint categories (n = 398)	Percentage
Coughing	13.6%
Throat complaints	8.0%
Ear complaints	5.8%
Skin complaints	5.5%
Eye complaints	5.5%
Level of urgency	
U2	37.9%
U3	41.2%
U4	8.8%
U5	12.1%
Week/weekend	
Week	20.6%
Weekend	79.4%
Compliance	
Yes	67.3%
No	32.7%

Distribution of triage characteristics in percentages.

3.4. Compliance

Observation of patient behavior post-triage found 67.3% of patients complying with the advice given. Compliance among those who received a U5 level of urgency and were advised not to contact a doctor was higher, with 77% of patients refraining from GPC utilization (Table 3). Compliance was associated to triage completion in the weekend (OR 2.9, 95% CI 1.3 – 3.9, $p<0.002$) versus triages conducted on weeknights. Other examined factors such as age, sex and triage urgency were not found to impact compliance significantly in this study.

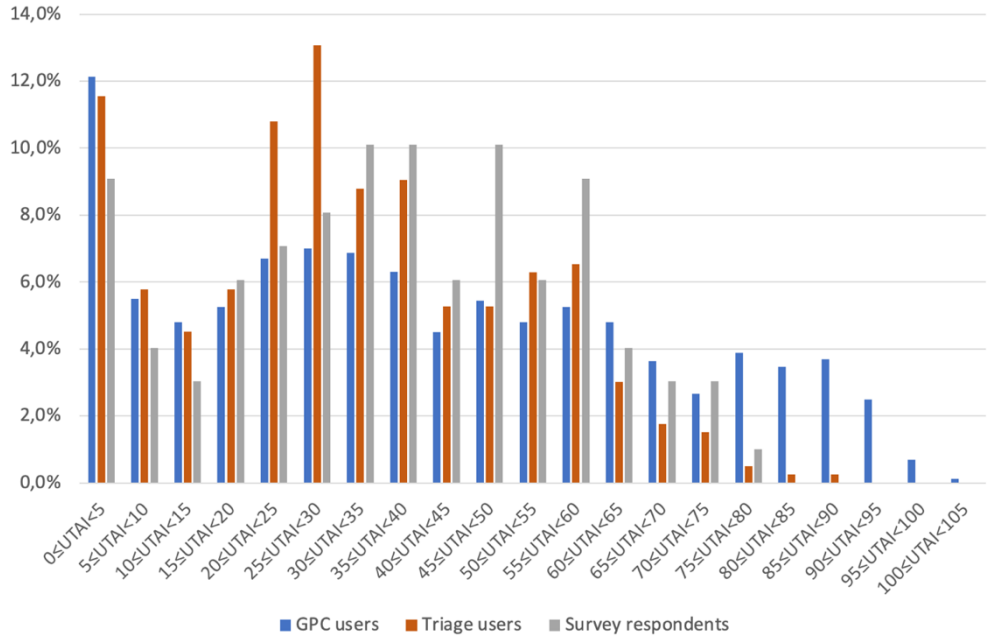
There is a marked discrepancy with the reported intention to comply, to which 88.9% of participants answered affirmatively. Moreover, when comparing actual behavior to the stated intention, 73.2% of cases showed to correspond, resulting in 26.8% of participants acting contrary to what they indicated (Table 4).

Table 4. Questionnaire data

Completed questionnaires (n = 99)	Percentage
Question 1	
Yes	97.0%
No	3.0%
Question 2	
Yes	88.9%
No	11.1%
Question 3	
Other	3%
I am reassured	2.0%
I am not yet reassured	5.1%
I am adequately informed	15.2%
I have not been able to explain my situation	5.1%
I trust the advice	69.7%
Question 4	
Yes	12.1%
No	87.9%
Question 5	
Yes	70.7%
No	29.3%
Compliance <> Intention to comply	
Yes	73.2%
No	26.8%

Distribution of questionnaire results in percentages.

Figure 2. Age distribution per user group



Distribution of age in percentages per examined user group. This figure illustrates the relative proportions within each study group and presents them side by side with the other groups to visualize the degree of comparability.

A large fraction of survey respondents said they trusted the advice (69.7%) or felt adequately informed (15.2%), which was a driving factor in their intention to comply with the advice. In contrast, people who disagreed with the advice mainly indicated that this stemmed from a feeling of insufficient reassurance (5.1%) or the experience that they had not been able to tell their story sufficiently (5.1%). The latter is confirmed by the responses to question 5, of which 29.3% indicated that they were not able to adequately represent their issues using the triage protocol.

People used the tool almost exclusively with the preconception of requiring a doctor (97%). 12.1% subsequently indicated that the generated advice modified their health care seeking behavior. (Table 4)

4. Discussion

4.1. user demographics

The observed distribution in age of triage users echoes the findings of previous studies[1, 2, 6, 18, 19]. The userbase is skewed more toward younger subpopulations[1, 2, 11, 21], with a distinct peak around 20-30 years of age. A second outlier is seen in children under 5 years of age. This perceived need for care is reflected in the GPC usage statistics and fits within a previously observed considerable unplannable care demand among young children[15], at times straining the capacity of the available care services.

The high prevalence of elderly people using the out-of-hours care services is not reflected in triage utilization. This discrepancy is presumably related to factors like digital literacy, accessibility of internet, and communication habits, and is in line with current evidence[2, 6, 13, 19]. The implementation of such systems could provide younger subpopulations with a low threshold pathway into the health care system but should be implemented within a range of complementary and varied methods of contact[9, 13, 21] to ensure a similar accessibility of care to the general population[1, 2, 22].

Important to note is that a proportion of completed triages is performed by a proxy, which is illustrated by the triage results of young children. This evolution could continue to other age groups as well. To objectify this, future surveys could differentiate between the age of the respondent and of the patient.

The majority of triage users were female, a finding that is consistent with current evidence[13, 19]. Interestingly, this female predominance is also observed in the GPC users, potentially indicating that women seek out-of-hours care more frequently.

4.2. Triage characteristics

A high incidence of complaints related to infectious diseases was observed. It is important to note that data collection was limited to the month of march, during which infectious disease burden on the health care system generally remains considerable.

A comparatively high number of cases are triaged as high acuity, with U1 and U2 recommendations in particular exceeding the expected incidence in primary care. This is presumably due to the risk averse predisposition of most digital triage algorithms[18, 23, 24]. Evidence suggests this could negatively impact compliance due to the perception that the advice is excessive and inappropriate for the symptoms experienced[1, 2, 25].

4.3. Compliance

An overall compliance rate of 67.3% appears to be in line with previous studies[1, 2, 19, 25]. It is important to note that this is a measurement of mapped behavior post-triage as opposed to gaging intention to comply. Comparison of the collected data on these outcomes reveals a significant discrepancy, with 88.9% reporting planned adherence to the advice given and a subsequent mismatch of 26.8% between resolution and action. This disparity evokes reservations about using stated intent as a representative outcome for behavior and highlights the added value of objectifiable behavioral mapping performed in the intended care setting.

Interestingly, compliance among participants receiving a self-care advice was higher still (77%), even though this most often differed from their original intention, considering 97% of triage users thought they needed a doctor prior to triage. This finding potentially illustrates the algorithms' ability to modify health care seeking behavior. Additional research on potential influencing factors, such as type of complaint and the comprehensiveness of provided advice would be of added value. Moreover, it is important to examine compliance of users receiving U1 advice.

Compliance was shown to be associated with triage during the weekend as opposed to during weeknights. A potential underlying driver contributing to this correlation could be a discrepancy between compliance during the day versus the night. Patients may perceive a higher threshold to contact the GPC during nighttime, postponing action until the next morning. There are arguments present in the studied datasets supporting this hypothesis, with potential matches missed due to the currently applied criteria defining compliance in this study, which requires patients to interact with the GPC within 3 hours of triage. These potentially missed patient flows could cause an underestimation of compliance during the night. Such mismatching would weigh more heavily on weeknight data due to the partial compensation by daytime figures on weekends. This potential confounder is certainly worth investigating, through analysis of daytime versus nighttime triage. Additionally, a possible redefinition of the data-matching criteria and identification of experienced barriers to seek contact could serve as a foundation to improve both the tool and study methodology.

3.4 Study limitations

There is a risk of selection bias in data collection from the triage algorithm, since U1 advice and dental complaints were excluded from the sample. Additionally, while a large and diverse GPC region was selected for this study, this sample alone might not be fully representative of the national context. This also applies to the duration of data acquisition, with multifactorial variations between different times of year potentially impacting findings. Patients had the choice to exclude their anonymous data from scientific research by actively opting out, which could have introduced limited non-response bias. There is no data on this, as all interaction with the tool was deleted.

The survey on intention to comply was prone to non-response bias, a common risk for this type of methodology. An effort was made to lower the participation threshold by directly integrating the survey into the triage advice page and limiting the number of questions, as well as displaying the progress through the survey to the patient. Comparison of demographic data of the triage tool and the survey participants revealed only minor differences, which we believe does not limit the transferability of findings significantly. Moreover, the questionnaire is inherently at risk of response bias, with self-reporting by participants potentially leading to inaccurate answers. This is especially relevant if the truth is considered sensitive or undesirable. In this case, admitting to defying the provided advice might be avoided by participants. This potential bias is in part the subject of RQ3 and is countered by correlating the answers of the questionnaire to the patient's observed behavior, thus quantifying this false reporting, and shedding light on the validity of intention to comply as a measurement for compliance.

4.5. Future research

There is a scientific need for a more extensive and multifaceted research effort, performed in the setting for which the specific tool is intended. One of the foundational aspects of such exhaustive exploration would consist of an in-depth analysis of demographic data of triage tools, the corresponding GPCs, and by extension the regional population. In doing so, a more nuanced perspective can be formed on non-plannable care utilization and the impact of subsequent management strategies on demand and accessibility of care.

In addition, this entails the expansion of data collection which enables mapping of seasonal evolution of complaint and disease incidence, potentially generating relevant data for disease monitoring at the macro-level.

A larger sample size could benefit research on compliance as well. This is especially true in case an increased variety in population-density, geographical distribution, and economic situation is achieved. A more accurate representation of the general population could improve transferability of findings. In addition, an increase in statistical power and a more comprehensive analysis of potential influenced factors could prove scientifically relevant.

References

1. Pairon, A., H. Philips, and V. Verhoeven, *A scoping review on the use and usefulness of online symptom checkers and triage systems: How to proceed?* Front Med (Lausanne), 2022. **9**: p. 1040926.
2. Turner, J., et al., *Health Services and Delivery Research*, in *Impact of NHS 111 Online on the NHS 111 telephone service and urgent care system: a mixed-methods study*. 2021, NIHR Journals Library Copyright © 2021 Turner et al.: Southampton (UK).
3. Aboueid, S., et al., *Young Adults' Perspectives on the Use of Symptom Checkers for Self-Triage and Self-Diagnosis: Qualitative Study.* JMIR Public Health Surveill, 2021. **7**(1): p. e22637.
4. Dickson, S.J., et al., *Agreement and validity of electronic patient self-triage (eTriage) with nurse triage in two UK emergency departments: a retrospective study.* Eur J Emerg Med, 2022. **29**(1): p. 49-55.
5. Chan, F., et al., *Performance of a new symptom checker in patient triage: Canadian cohort study.* PLoS One, 2021. **16**(12): p. e0260696.
6. Chambers, D., et al., *Digital and online symptom checkers and health assessment/triage services for urgent health problems: systematic review.* BMJ Open, 2019. **9**(8): p. e027743.
7. Millenson, M.L., et al., *Beyond Dr. Google: the evidence on consumer-facing digital tools for diagnosis.* Diagnosis, 2018. **5**(3): p. 95-105.
8. Cross, S., et al. *Search engines vs. symptom checkers: A comparison of their effectiveness for online health advice*. in *Proceedings of the Web Conference 2021*. 2021.
9. Schmieding, M.L., et al., *Benchmarking Triage Capability of Symptom Checkers Against That of Medical Laypersons: Survey Study.* J Med Internet Res, 2021. **23**(3): p. e24475.
10. Ceney, A., et al., *Accuracy of online symptom checkers and the potential impact on service utilisation.* PLoS One, 2021. **16**(7): p. e0254088.
11. Sutham, K., P. Khuwuthyakorn, and O. Thinnukool, *Thailand medical mobile application for patients triage base on criteria based dispatch protocol.* BMC Med Inform Decis Mak, 2020. **20**(1): p. 66.
12. Tsai, C.-H., et al. *Exploring and promoting diagnostic transparency and explainability in online symptom checkers*. in *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*. 2021.
13. Arksey, H. and L. O'Malley, *Scoping studies: towards a methodological framework.* International Journal of Social Research Methodology, 2005. **8**(1): p. 19-32.
14. Schmieding, M.L., et al., *Triage Accuracy of Symptom Checker Apps: 5-Year Follow-up Evaluation.* J Med Internet Res, 2022. **24**(5): p. e31810.
15. Giesen, M.-J., et al., *The impact of demand management strategies on parents' decision-making for out-of-hours primary care: findings from a survey in The Netherlands.* BMJ Open, 2017. **7**(5): p. e014605.
16. Donovan, E., et al., *Digital interventions for parents of acutely ill children and their treatment-seeking behaviour: a systematic review.* British Journal of General Practice, 2020. **70**(692): p. e172-e178.
17. Elliot, A.J., et al., *Internet-based remote health self-checker symptom data as an adjuvant to a national syndromic surveillance system.* Epidemiol Infect, 2015. **143**(16): p. 3416-22.
18. Aboueid, S., et al., *The Use of Artificially Intelligent Self-Diagnosing Digital Platforms by the General Public: Scoping Review.* JMIR Med Inform, 2019. **7**(2): p. e13445.
19. Verzantvoort, N.C.M., et al., *Self-triage for acute primary care via a smartphone application: Practical, safe and efficient?* PLoS One, 2018. **13**(6): p. e0199284.
20. Meyer, A.N.D., et al., *Patient Perspectives on the Usefulness of an Artificial Intelligence–Assisted Symptom Checker: Cross-Sectional Survey Study.* J Med Internet Res, 2020. **22**(1): p. e14679.
21. Yu, S.W.Y., et al., *Triage accuracy of online symptom checkers for Accident and Emergency Department patients.* Hong Kong Journal of Emergency Medicine, 2020. **27**(4): p. 217-222.
22. Hill, M.G., M. Sim, and B. Mills, *The quality of diagnosis and triage advice provided by free online symptom checkers and apps in Australia.* Med J Aust, 2020. **212**(11): p. 514-519.
23. Gottliebsen, K. and G. Petersson, *Limited evidence of benefits of patient operated intelligent primary care triage tools: findings of a literature review.* BMJ Health Care Inform, 2020. **27**(1).
24. Semigran, H.L., et al., *Evaluation of symptom checkers for self diagnosis and triage: audit study.* Bmj, 2015. **351**: p. h3480.
25. Nijland, N., et al., *Patient use and compliance with medical advice delivered by a web-based triage system in primary care.* J Telemed Telecare, 2010. **16**(1): p. 8-11.

Appendix 1. Approval of the ethics committee of UZA-UAntwerp

ETHISCH COMITE

Prof. Dr. Peter Michielsen

Secretariaat

03 821 38 97

Professor Veronique Verhoeven

Gouverneur Kinsbergencentrum

Doornstraat 331

2610 Wilrijk

Dr. Anthony Pairon

FAMPOP

datum: 13/12/2021

Project ID 2021-0665 - EDGE n/a - BUN/EudraCT n/a

Project titel: Digital self-triage for urgent primary care: A quantitative evaluation of safety, accuracy, compliance and impact through data matching with iCAREdata.

DEFINITIEF GUNSTIG ADVIES

Geachte collega

Het Ethisch Comité van het Universitair Ziekenhuis Antwerpen en de universiteit Antwerpen heeft bovenvermelde studie besproken en geeft een gunstig advies dd. 13/12/2021.

De volgende bijlagen werden volgens de ICH-GCP richtlijnen door het Ethisch Comité goedgekeurd:

Document Type	File Name	Date	Version
GCP	Veronique-Verhoeven-25-06-2021-165210000309	25/06/2021	3.0
CV	Curriculum-vitae_EC_AP_v1	31/10/2021	1.1
Accompanying letter	AccompanyingLetter_EC_AP_v1-2	07/11/2021	1.2
Protocol	Study protocol MINDD_EC_AP_v1-2	07/11/2021	1.2
Insurance	Insurance_EC_AP_v1-2	07/11/2021	1.2
GCP	GCP_EC_AP_v1	12/11/2021	1
CV	UA_EC_WETL_Veronique_Verhoeven	18/11/2021	1

De volgende opmerking werd nog gemaakt:

Gelieve de bewaringstermijn aan te geven = 20 jaar.

Deze goedkeuring is geldig tot een jaar na bovenvermelde datum. Wij verzoeken u ons te melden wanneer de eerste deelnemer werd geïncludeerd, wanneer en waarom de studie (vroegtijdig) werd stopgezet of nooit werd opgestart.

Indien de studie nog loopt na een jaar verwachten we een follow-up rapport waarin eventuele voorvallen worden gemeld.

Tot slot wijzen we er op dat, voor in het UZA lopende studies, de ernstige ongewenste voorvallen dienen gerapporteerd via het incidentenmeldingssysteem.

Met vriendelijke groeten



Prof. dr. Peter Michielsen

Voorzitter Ethisch Comité UZA/UAntwerpen

Meeting Attendee Full Name	Meeting Attendee Qualifications
Ms Bettina Blaumeiser	Physicians
Emeritus Professor Hilde Bortier	MD, PhD
Prof. Dr. Patrick Cras	Vice-Chair, Physicians
Ms Ingrid De Meester	Pharmacologist
Ms Elyne Debaetselier	Nurse
Ms Lina Fierens	Nurse
Ms Kristina Granaas	Staffmember Research Dept. UAntwerpen
Ms Kristien Hens	Philosopher
Mr Kris Ides	Physiotherapist
Prof. Dr. Greet Ieven	Vice-Chair, Physician
Mr. Daan Kenis	Pharmacist
Ms Johanna Kwakkel-van Erp	Physicians
Dr Leon Luyten	Physicians
Ms Barbara Michiels	General Practitioner
Mr Peter Michielsen	Chair, Physicians
Ms Inge Michielsens	Legal
Mr Pieter Moons	Coordinator Bio- and Human Tissue bank
Mr Bernard Paelinck	Physicians
Mr. Michaël Pössel	Healthy volunteer phase 1
Ms Veerle Schoeters	Nurse
Mr. Kris Smulders	Nurse
Mr Jan Van den Brande	Physicians
Mr Filip Van den Eede	Physicians
Mr Pieter Van Dyck	Physicians
Mr Guy Van Honste	Patient Representative
Ms Dominique Van Praag	Psychologist
Mr. Yves Vandermotten	Healthy volunteer phase 1
Mr Thierry Vansweevelt	Legal
Ms Griet Wyers	Pharmacist

Universiteit Antwerpen
Faculteit Geneeskunde en Gezondheidswetenschappen

Family Medicine and Population Health

Evaluation of a digital self-triage system for urgent primary care through data matching with iCAREdata - Study protocol

Student: Anthony Pairon
Promotor: prof. dr. Veronique Verhoeven
Co-promotor: prof. dr. Hilde Philips
07/11/2021

Introduction:

The system for out-of-hours medical care in Belgium, consisting of general practitioner cooperatives and emergency departments acting as complimentary services, faces an increasing workload, burdening healthcare providers and stretching budgets.

The increase in demand is caused in part by a growing and ageing population. Another important factor however is the inappropriate use of the available services, because patients are required to assess the urgency of their medical complaints and decide on the appropriate caregiver without guidance. This contrasts with the emphasis of modern medicine on the informed patient as an integral aspect of personal health management.

Self-triage systems could assist patients in their healthcare seeking behaviour, helping them to get the right care at the right time and improving the efficiency of the available care services.

Online triage tools are currently available in several countries, such as NHS24 as a part of the NHS service in the UK, moetiknaardedokter.nl in the Netherlands, which is advocated by several GPC OOH services, and SNS24 in Portugal. Research supporting these digital assessment services for patients seeking guidance about urgent care is largely qualitative and focuses on patient satisfaction and user friendliness. Data on the accuracy of these tools and the potential impact on the usage of OOH medical services is scarce and mainly obtained from small study samples over short periods of time.

In general, algorithm-based triage tends to be more risk averse than the triage of health professionals. Two components are important in the context of accuracy of these triage tools: (1) are they safe, i.e., is the amount of under-triage limited, and (2) are they efficient, i.e., is the proportion of over-triage acceptable. The limited reports available seem to indicate a favourable safety outcome.

Lastly, there is little knowledge about the possible impact of these tools on use of OOH medical services. Three reports indicate, with low levels of evidence, that they may lead to a reduced number of consultations.

ICAREdata, which complies with the most recent legal, ethical and privacy aspects required in Belgium and Europe, contains (almost) real-time data of patients using OOH care in a large part of Flanders. The system offers a unique opportunity to study safety and accuracy aspects of the online triage tool by using routine registration data on a large scale. Thus, this research has the potential to contribute significantly to the validity of these services, further empowering patients in their personal health management.

Study objectives

This project will evaluate an existing digital self-triage tool aimed at assisting patients in their out-of-hours healthcare seeking behaviour. The triage system, moetiknaardedokter.be, is based on the Dutch triage standard and will be available in four well-defined regions through the websites of the corresponding general practice cooperatives (GPC) as well as via a free app for mobile phone.

These regions are included in the iCAREdata project of Antwerp University, which registers every out-of-hours (OOH) consultation at the GPC and ED. Linking of both databases allows for the evaluation of patient flow, enabling us to answer several research questions.

Research Questions:

RQ1 - Demographic: How often is the triage tool used over the study time, by which patients (age and sex) and for which complaints?
RQ2 - Compliance: What is the compliance with the given triage advice and how is it influenced by the nature of the complaint and other influencing factors?
RQ3 – Accuracy & safety: How correct was the triage advice for people consulting a GPC or ED afterwards, based on concordance or disconcordance with the expert opinion of the GP/emergency doctor?
RQ4 - Impact: What is the evolution of the OOH-use in the studied region during the study period?

A second, qualitative chapter of the project will focus on user-friendliness and patient satisfaction.

Method

Study design:

Retrospective epidemiological study

Study population:

The patients attending GPC's in four well-defined regions over the study period. The studied area encompasses Antwerpen-Oost, Zuiderkempen, Noordrand and Turnhout, covering approximately 759000 inhabitants.

Procedure:

The online triage tool moetiknaardedokter.be is available through the website of the participating GPC's. Additionally, patients have the option to use the mobile app, which is CE-certified, thus conforming with applicable EU (safety) regulations. Routine registration data of this tool (age, sex, medical complaint and the generated advice) will be delivered to us free of charge by the Dutch ICT company which supports the system.
A second dataset will be extracted from ICAREdata, containing anonymous data of every patient using OOH care, both GPC's and emergency departments, in the study regions.
Our team will link both datasets using the routine registration data. The data matching will be automated, based on three variables: year of birth, sex and date of contact. In case of doubles, a researcher with an MD will manually examine the individual matches and include the medical complaint as an additional variable to find the correct match.

Exclusion criteria:
Matches are excluded when the time between digital triage and first contact with a healthcare provider exceeds two hours.
If it remains unclear which match is correct, both will be excluded.

Analysis
Descriptive statistics will be performed two answer research question 1, 2 and 4. Furthermore, the amount of undertriage (a marker for safety) and overtriage (a marker for efficiency) will be calculated.

Study administration

Data collection and management

The implementation of the tool on the website of the GPC is the responsibility of the GPC itself. However, in the context of our collaboration, we obtained legal advice concerning medical and GDPR-related liability. This resulted in small changes to the website: a clear "instructions for use", a statement regarding the scientific basis of the tool – which was already present on the website -, adding a contact person and a statement on how the data will be used.

Data collection and management will be done within the legal, ethical and privacy related framework of iCAREdata. The data delivered to us is strictly anonymous, since personal data such as name and address are not registered by the triage tool. iCAREdata registers the social security number, but this is coded by a third party and is not delivered in the output data. No extra data will be generated in the context of our study. Thus, this guaranteed anonymity obviates the need for an informed consent. However, users of the tool will be able to opt out from the fact that their anonymous data will be used for scientific research by ticking a box. No sample is taken and thus no sample size is calculated. All data registered during the study period of 1 year will be used.

Ethical justification

All data used in the study is routinely registered in the context of both systems and strictly anonymous, making it impossible to trace the accumulated information to any individual participant. No additional data is collected.

Proposed timeline

- January 2021 – September 2021:
  - Recruitment of GPC's and implementation of the triage tool.
- September 2021 - August 2022:
  - Proposed period of collected data~~ collection~~ that will be used.
- November 2021:
  - Application to the UZA ethical committee.
- January 2022 – September 2022:
  - Retrospective data extraction, matching and analysis.
- October 2022 – January 2023:
  - Writing the thesis.
- February 2023:
  - Proofreading and integration of feedback in the final text.
- May 2023:
  - Submission master thesis.
- May/June 2023:
  - Presentation and defence of master thesis with the aim of obtaining the degree of Advanced master in Family Medicine.