UNIVERSITÉ DE GENÈVE
FACULTÉ DE MÉDECINE

Clinical Medicine Section
Department of Primary Care

INSTRUMENTS FOR OPTIMIZING PATIENT CARE IN THE OUTPATIENT EMERGENCY UNIT, WITH A FOCUS ON PERCEPTION OF CARE

Thesis submitted to the Faculty of Medicine of

the University of Geneva

for the degree of Privat-Docent

by

Dr. Hervé SPECHBACH

(*Geneva*)

(2024)

Contents

Summary

Overcrowding in emergency departments (EDs) is a common problem worldwide, often resulting in long wait times, decreased quality of care, patient dissatisfaction with the service, and increased costs. The impact on caregivers is also significant, with an increased risk of burn-out, absenteeism, and high staff turnover. The two leading contributors of overcrowding at our ED are the input (volume of patients arriving) and throughput (time to process and/or treat patients). The main causes of input overcrowding are non-urgent cases, frequent users, and the poor and uninsured who lack primary care, mainly migrants.

We first identified the time spent waiting to see a doctor as one of the most important variables associated with patient satisfaction in the ED. Thus, we investigated patients' perception of wait time and developed a novel tool permitting to change the perception of wait time, with the hypothesis that it would impact on satisfaction.

As communication is a key factor for patient satisfaction and better adherence to treatment, we then took into account the fact that 10% of our patients do not share a common language with caregivers in the outpatient emergency patient of Geneva University Hospitals (HUG), as well as those with intellectual deficiencies and deaf persons. This led to the development of novel communication methods, such as BabelDr, a speech-enabled fixed-phrase translator for emergency settings, and its extension, PictoDr, which allows communication exclusively through pictographs. Several other innovative tools are currently under development, such as the use of avatars for the deaf.

This privat docent thesis represents the results of the author's research, which focuses on instruments to optimize patient care and enhance patient and caregiver satisfaction in the outpatient emergency unit.

Abbreviations

ED	Emergency Department
NHS	National Health System
HUG	Geneva University Hospitals
SMS	Short Message System
OR	odds ratio
MT	machine translation
GT	Google Translate
ARASAAC	Aragonese Center of Augmentative and Alternative Communication
AI	artificial intelligence
UMLS	Unified Medical Language System
CUREG	University of Geneva Commission for Ethical Research
SGAIM	Swiss General Internal Medicine Association

1. Introduction

1.1 Patient satisfaction in the emergency department (ED) in the context of overcrowding

Similar to other hospitals in Switzerland and many countries worldwide, Geneva University Hospitals (HUG) are facing the challenge of overcrowding in the ED. First described in the 1980s, this has now become a chronic situation, significantly intensifying over the last few years (1). This worrisome trend has an impact on the quality of care, patients' perception of care, patient satisfaction, and costs. It also increases medicolegal risks. The impact on caregivers is high, with an increased risk of burn-out. In the current situation of a worldwide lack of caregivers, absenteeism is an additional concern (2).

When considering the aspect of patient (or "consumer") satisfaction with healthcare provision, political interest has increased in the past 20 years, notably in the United Kingdom with the advent of the *Patients' Charter* of 1991 and the *NHS Plan* in 2000. The latter laid out the government's intention to increase funding (investment) and reform an ageing health system "fit for the 21st century" and more adapted to patient expectations (3). A similar trend has been observed in other developed countries (4). In parallel, hospital administrators have placed a greater emphasis on the satisfaction of patients consulting the ED and it is now considered as an important quality indicator (5). The ED is the hospital's showcase and administrators are well aware that in a service-oriented healthcare market, the "customer" experience is a key issue in this sector (4). Additionally, in the context of Internet and the risk of a wide and rapid dissemination of negative ratings, the word-of-mouth market is powerful and has been emphasized by politicians since many years (6).

However, the level of patient satisfaction in the ED is not only a commercial issue reflecting on the image of the hospital, it has also an impact on the quality and continuity of care. For example, Dubina et al reported that satisfied patient-clients will be more adherent to their treatment (2). Furthermore, it is well known that a patient satisfied with the stay in the ED will be more inclined to return to the same hospital for specialized care (7).

HUG is a 2000-bed primary and tertiary urban teaching hospital serving the canton of Geneva and surrounding population (approximately 800,000 inhabitants). The HUG adult outpatient emergency unit received approximately 42,000 patients in 2023, of whom 7% were hospitalized. All patients admitted are triaged using the four-level Swiss Emergency Triage Scale. More than 90% of patients are triaged according to level 3, i.e., semi-urgent conditions requiring medical evaluation by a doctor within 120 minutes. However, it has been reported that patients are at risk of leaving without having seen a doctor after 1 hour of waiting time with the known risk of a return visit to the ED in the next few days (8). Moreover, the waiting time (in the ED waiting room) before seeing a doctor is perceived as the worst part of the patient journey (4). The complexity of patient satisfaction with the HUG outpatient emergency unit is also related to the fact that many patients come from a great diversity of cultural backgrounds, including the fact that we do not share a common language with around 10% of our patients, despite our multilingual staff.

Paying attention to the satisfaction of patients is important in an emergency outpatient unit that also welcomes visitors with non-urgent conditions, particularly with regards to the "point of view paradox" in that the less severe the illness, the greater the expectations of patients regarding non-clinical service factors (9). It is also reported that people with a low emergency triage level are more likely to leave without being seen. Moreover, in a country like Switzerland where patients pay increasingly a greater part of non-reimbursed health costs, a certain level of service is expected.

1.2 Characteristics of overcrowding in the ED with a focus on an outpatient emergency unit

The standard to describe overcrowding in the ED (10) is to divide the problematic into three specific periods: the *input* (the volume of patients arriving); the *throughput* (the time to process and/or treat patients); and the *output* (the volume of patients leaving the emergency department (boarding time). In the specific situation of the HUG emergency outpatient unit, the output is not the most relevant level of action as we hospitalize only 7% of patients. The two leading contributors of overcrowding are the input and throughput. Moreover, since 2019, we have observed an increased input of 30%, which represents 100% since 2015. The main causes of input overcrowding are non-urgent cases, frequent users, and the poor and uninsured who lack primary care, mostly migrants. We are principally confronted here by the malfunctioning of healthcare services in the community (11) and thus we are unable to influence the input. Regarding migrants consulting the outpatient emergency unit, the major need was to have a reliable tool permitting to interact with them 24h/24h.

1.3 Possible actions to meet the challenges of ED overcrowding and patient dissatisfaction/satisfaction in an outpatient emergency unit

Healthcare professionals working in the ED can act on several levels. First, they can describe the situation and alert health system stakeholders (12). Another level is to inform the population about illnesses to improve their health literacy, a strategy known to help patients to be more compliant with treatment and (theoretically, at least) decrease return visits to the ED. However, the effect of this action is controversial in the literature as a means to avoid return visits (13-14). For ED caregivers, the input is not the easy way to act to mitigate overcrowding as it needs a political input and more general actions (15).

An additional level of action is to act on the processes that can be controlled during a stay in the ED (*throughput*) (10) before, during and after seeing the doctor, including the quality of discharge (*output*). The actions here are in the hands of the caregivers as the process takes place in the unit. Among other aspects, this can be achieved by developing instruments to optimize patient care, with a focus on the perception of care by the patient. Here, I focus on the time before seeing a doctor, a part of the throughput. As people come to the ED to see a medical doctor, the least pleasant waiting time spent in the ED is the time before actually seeing a doctor (4). Additionally, it has already been reported that a short waiting time (particularly if shorter than expected) is a key factor for recommending the ED (6). However, the wait time before seeing a doctor fluctuates throughout the day according to peaks of activity that are currently not yet predictable in the HUG setting, similar to most health facilities worldwide.

1.4 Satisfaction in the ED with a focus on the wait time perception before seeing a doctor

Patient satisfaction in the ED is a holistic experience, which is not easy to understand as several primary implications must be considered, i.e., cultural aspects, language barriers and psychological behavior (16). Many factors influence satisfaction (6), such as cleanliness, caregivers' appearance, the communication skills and competences of caregivers, accessibility, privacy, security, and billing, but mainly waiting time and information delivery (4). I emphasize here the waiting time before seeing a doctor as it is one of the main causes of satisfaction/dissatisfaction (4). Notably, the relationship between the perceived and expected waiting time appears to be a key factor (3) .

In the context of patient satisfaction, we are faced with the dilemma of patient (client) expectations and the perception of care during the ED stay, including waiting time. In 1985, David Maister published a model to analyze the psychological aspect of patient satisfaction in the ED (5): *satisfaction* = *perception - expectation.* In addition, there is a causal link between satisfaction and adherence to the prescribed treatment, the number of complaints, the public opinion of the hospital in general, and the satisfaction of staff working in the ED.

In their systematic review, Taylor et al (3) found that the most frequent factor influencing the overall satisfaction of patients was the perceived wait relative to the expected wait, with a strong correlation with overall satisfaction ($p<0.001$). Information and perceived wait (but not actual wait) correlated with global satisfaction ($p<0.001$), as well as information about the level of emergency triage ($p<0.01$) (7) and, in a minor way, staff attitude ($p<0.05$) (17), the actual wait to see the doctor ($p<0.003$) (18), and total time spent in the ED ($p<0.01$) (18) (6). As waiting time seems to stop when the patients see the doctor, one of the most important variables associated with satisfaction in the ED is therefore the time spent waiting to see a doctor (6).

Concerning this specific time spent before seeing a doctor, in the **first article** described here (19), we addressed the issue of the specific time spent before seeing the doctor and impacting in consequence on overall satisfaction was correct triage (odds ratio [OR], 5.37 [95% confidence interval [CI], 2.46-11.73]; p<0.001). In decreasing order of importance, this was followed by the “feeling of being forgotten” (OR, 3 [95% CI, 1.47-6.15]; p<0.001) (the question was “I didn’t feel forgotten”), non-respect of privacy (OR, 0.21 [95% CI, 0.05-0.88]; p=0.03), and the actual waiting time (OR, 0.98 [95% CI, 0.98-0.99]; p=0.03).

Following this study, we developed a Semiautomatic Text Message (SMS) system to allow patients to wait outside the outpatient emergency unit, which is available in eight languages (Arabic, English, Farsi, French, German, Italian, Spanish and Tigrigna). This system is an extension of a similar project developed in the pediatric ED at HUG (20). Patient satisfaction with the use of the SMS system was analyzed in the **second article** presented here (21). Notably, we observed that the waiting time outside the emergency outpatient unit does not impact on overall satisfaction and more than 95% of patients were very satisfied with their waiting time before seeing a doctor.

To the best of our knowledge, no similar studies have been reported in the literature, apart from Puthenparampril (22) who showed that the use of text messaging to mitigate the waiting time before surgery enhanced patient satisfaction. Waiting outside the ED allows patients to choose the place where to wait, with the comfort and privacy they require. This option has many advantages. It allows the patient to go to one of the rest areas proposed by the hospital (e.g., cafeteria, adjacent green spaces, meditation room) and has a positive impact on the perception of the waiting time (23). It also enhances patient safety by decreasing the risk of contracting an airborne infection (24) by reducing overcrowding in the ED waiting room, which also negatively affects the patient and staff experience.

1.5. How to have an impact on the satisfaction of patients who do not share a common language with caregivers in the ED?

We cannot envisage actions on ED overcrowding and the associated development of instruments to optimize the overall satisfaction with the perception of care without considering the fact that 10% of our patients do not share a common language with caregivers (HUG, internal data). The international city of Geneva lies at the crossroads of Europe with a long history of welcoming refugees. It also hosts the highest number of international organizations and diplomatic missions in the world. Thus, the language barrier is encountered daily in our ED, and it was obvious that we should take it into account as a significant factor impacting on patient satisfaction and safety. Moreover, as described by Sartini

(11), one of the major causes of ED overcrowding is the lack of primary care access, particularly for migrants.

In the context of the current migration crisis and the next potential, climate-related, migration emergency, we will be increasingly faced with this problematic of the language barrier and the needs of a multicultural healthcare system. These two elements complicate the quality of communication between the patient and caregiver in order to ensure that the patient is satisfied with the stay and care given in the ED. Communication is a key factor for patient satisfaction, perception of care, and better adherence to treatment. We have also observed the same difficulties with people with intellectual deficiencies (25) or of low socioeconomic status and it has become increasingly evident that we need to develop augmentative and alternative communication in the ED, i.e., communication methods used to supplement/replace speech or writing for those with impairments in the production or comprehension of spoken or written language. The quality of explanations and answers to questions need to be correlated to patient satisfaction, including pain or stress perceptions (26). In this context, we need to develop tools to communicate with such populations in the context of patient-centered care, with a related specific communication at each touch point with patients, as described recently in the *New England Journal of Medicine* round table entitled “What really matters to patients” (27).

1.6 Optimizing patient care and satisfaction when patients do not share a common language with the caregiver.

Information received from caregivers and medical doctors and the possibility to be understood is a key factor of quality of care, morbi-mortality and patient satisfaction (28). In the ED, the triage process with questions understood by patients is crucial and can increase psychological distress and dissatisfaction if this is not the case (29). As mentioned, in the context of the current refugee crisis (or a future similar emergency), our ED unit is regularly confronted with patients with whom we do not share a common language. Indeed, a recent Swiss study confirmed the difficulties encountered in this area and stated that 90% of medical doctors face a language barrier at least once a year, with 30% reporting that they have to deal with it once a week (30). Another analysis conducted in a Swiss French-speaking ED found that language barriers were present at least once a week for 71% of doctors interviewed (spring congress of SSMIG, Basel 2024). In the ED at HUG, 52% of patients are non-Swiss and 10% speak no French at all (HUG internal data). Notably, the ED serves as the entry point to the healthcare system for such a population. This represents an ethical and safety problem for the patient, but also with a very expensive impact on healthcare costs. For example, an estimated US$ 73 billion are spent annually in the USA as a consequence of communication problems in healthcare (31). The language barrier in healthcare is a worldwide problem, with an important impact on the diagnosis, treatment and adherence to treatment due to miscommunication and/or absence of communication.

The preferred solutions to mitigate medical errors related to the language barrier are professional interpreters (if available), the only persons who are able to include cultural subtleties in verbal exchanges, then ad hoc interpreters (relatives or caregivers who can translate), and machine translation (MT) as a last solution (32). For professional interpreters, Flores et al (31) suggested the need for 100 hours' training to have an impact on reducing errors and to improve quality of care and patient safety (32). Such a solution is expensive and not reimbursed by Swiss health insurance companies. As an example, for the HUG Department of Primary Care, 21,000 hours are billed annually by professional interpreters and 74,000 for the entire HUG.

A diversity of apps have been developed in the field of translation (33), both fixed-phrase translators and general MT apps, which were first rule- or statistical-based, rather than deep learning-based (neural). Both types have their strengths and limitations. For the fixed-phrase translator, the advantage is the accuracy of translation, integrating the specific context, but the disadvantage is that the sentences are limited and the relation patient-caregiver is not very natural and without vocal recognition. For MT apps, the advantage is vocal recognition, but there is a risk of error, mainly in the medical context, as they are not designed for this situation. The change from statistical to neural MT has improved the quality of translation, but in the field of medical interviews, the results are not always safe and satisfactory, depending on the language (34).

Concerning MT apps, such as Google Translate, the advantage is that they are a low-cost solution and always available, but they are not developed for medical use. In 2016, the study by Flores et al (32) evaluated the accuracy of Google Translate for educational material for diabetes and showed that its accuracy decreased according to the complexity of the sentence and depending on the language. However, such a study was conducted before the integration of neural systems and no longer reflects the actual situation. More recently, in 2023, 40 English recordings translated into Spanish and 10 from Spanish to English in the field of pediatric neurosurgery demonstrated a high level of accuracy with Google Translate after evaluation by a medical interpreter (35). This shows that such a system is improving, but there is still a need to test other languages in a more complex setting to assess patient preferences, as well as the cultural context. There is also a need to compare different apps each other.

However, such a system like Google Translate is not allowed in Switzerland according to the Swiss Data Protection Law. At present, Microsoft Translator is allowed at HUG as the data appear to be protected and are not transmitted outside the European Union. A recent study concerning the use of Microsoft Translator was conducted in Geneva (34) and included 40 health professionals performing

52 consultations in 13 languages. This is a unique study evaluating voice-to-voice interaction between the caregiver and patients. Results showed that caregivers could reach their goal in 43/52 (82.7%) consultations but were satisfied with the communication in only 8/52 (53.2%) cases. Both patients and caregivers were understood in 34/52 (65.4%) consultations. Satisfaction decreased for languages outside the European Union due to poor vocal recognition.

For all the above-mentioned reasons and with the aim to have a safe, reliable and adaptative tool, we have developed the "homemade" Babeldr (www.babeldr.unige.ch) since 2015, a new type of speech-enabled, fixed-phrase, translation tool for medical dialogue in a collaborative venture between HUG and the University of Geneva Faculty of Translation and Interpreting. The project was funded by the HUG private foundation and it was awarded the Unitec prize (Office of Technology Transfer, University of Geneva) in 2015 and the HUG Innovation prize in 2018. The system is neither an MT that translates word-by-word with vocal recognition like Google Translate, DeepL or Microsoft Translator, nor a fixed phraselator like MediBabble that proposes pre-translated sentences. It is rather a system falling between these two categories. The system permits to speak using vocal recognition, then the best suitable recognized phrase is written (using both statistical and neural techniques) and the doctor chooses to translate them if he/she thinks that the sense is correct. The patients see the medical doctor speaking in the system and they can read and hear the sentences as many times as they want.

In BabelDr, multiple variations of vocal input with the same meaning will give the same sentence. For example, if the caregiver asks: "Avez-vous de la fièvre?", "Etes-vous fiévreux?", "Fièvre avez-vous?", BabelDr will propose the same sentence, i.e., "Do you have fever?", as well as pictograms with numbers, or the possibility for the patient to write something on the screen. We chose the pictographs ARASAAC (a Spanish pictograph repository: https://arasaac.org) after having tested them against SCLERA (an open content pictograph database: www.sclera.be). If we ask an open question, e.g., "Montrez-moi une position où vous avez plus mal?" ("Show me a position where you had the most pain?"), BabelDr proposes pictographs for answer by the patient (Figure 1). Conversely, when we ask : "Où avez-vous mal?" ("Where does it hurt?"), BabelDr will propose a closed question: "Pouvez-vous me montrer avec le doigt où est la douleur?" ("Can you show me with your finger where it hurts?") (Figure 2). We have 12,000 core sentences included with around one million variations, thus permitting to translate in the target language only the 12,000 core sentences for cost efficiency.

Figure 1: Example of Arabic output for the input: “Montrez-moi une position où vous avez plus mal?” (“Show me a position where you had most pain?”) where BabelDr proposes pictographs for answer by the patient. (Reproduced from https://babeldr.app.unige.ch/babeldrclient/#!/.)

Figure 2: Example of Arabic output for the input: “Où avez-vous mal?” (“Where does it hurt?”) where BabelDr proposes a closed sentence as it is a unidirectional system. (Reproduced from https://babeldr.app.unige.ch/babeldrclient/#!/.)

The patient sees the phrase written and can also hear it. There are only closed sentences and patients answer "yes", "no", or with pictographs. In the ED setting, the Swiss Emergency Triage Scale regroups hundreds of fixed closed sentences all included in the BabelDr tool. This point is important as our study showed that the major determinant of perceived waiting time was the appropriateness of having been classified in the correct degree of emergency according to the Swiss Emergency Scale (19). The translation of the sentences has been validated and verified by three professional interpreters of the University of Geneva Faculty of Interpreting and Translation. Moreover, the system allows us to choose the patient gender and the appropriate dialect for Arabic, which is not the case for example in Google Translate. The system is “homemade”, thus permitting us to adapt it according to the situation. For example, during the COVID-19 pandemic, we developed a specific domain in BabelDr with approximately 100 sentences needed at the HUG screening centre. We were also able to include the Ukrainian language in only one month at the beginning of the migration crisis to meet the needs of this population.

In 2017, we tested BabelDr versus Google Translate with a medical doctor at our hospital who was asked to diagnose standardized Arabic-speaking patients with abdominal pain, based on two scenarios (appendicitis and cholecystitis). We observed that BabelDr was more precise than Google Translate (OR, 0.04 [95% CI,0.02-0.12]; p<0.001) (36). We also tested our ergonomic design with the most used, fixed phraselator in the USA, MediBabble, regarding minor clicks (2.7 clicks [95% CI, 1.8-3.7]; p<0.001) and time for each interaction between doctor and patient (11 seconds [95% CI, 4.6-17.3]; p<0.001). Similarly, BabelDr was considered to be superior. Our hypothesis is that BabelDr has a speech recognition component, but not MediBabble. Thus, it is obvious that when we can vocalize in the system instead of typing text and searching with keywords, it is shorter and we need only minor clicks.

The **third article** (37) describes the characteristics of BabelDr when used in the laboratory of HUG with Arabic standardized patients and doctors from the ED, as well as general practitioners in the city of Geneva. We observed a high satisfaction with the tool, due to the added value of the voice recognition component, which permits to interact in a more natural way with the patient. All medical doctors concluded that they would appreciate the opportunity to use BabelDr on a daily basis with their own patients.

In the **fourth article** (38), we analyzed the satisfaction of both patients and caregivers with the use of BabelDr in the real life situation of the emergency outpatient unit. Again, the overall satisfaction was very high and all doctors were able to make a correct diagnosis and conclude with their patient. The voice recognition component was particularly described as a characteristic contributing to caregiver satisfaction.

With the aim to promote augmentative and alternative tools for communication in the ED, we have developed another tool since 2021, PictoDr, from the project PROPICTO (PROjecting spoken language into PICTOgraphs) a French-Swiss bilateral collaboration funded by the French National Research Agency and the Swiss National Science Foundation (https://propicto.demos.unige.ch/pictodrclient/). This tool allows communication to be conducted exclusively through pictographs and is targeted at a minimally verbal population (e.g., people with an intellectual deficiency, a cognitive disorder or limited health literacy). It also allows to cover the barrier gap for languages not yet included in BabelDr or for patients with a cognitive disorder. Individuals who use augmentative and alternative communication tools are particularly vulnerable when they need to go to the ED (39). The use of pictographs has been described as useful in this setting and can diminish stress, post-traumatic syndrome disorder or pain perception in the context of disaster or critical care patients transported by ambulance (26).

There are different systems permitting to generate pictographs, but they are generally very limited, such as “talking pictures”, while others like “PictoBert generate word-based pictographs with the risk of error in the medical domain. With PictoDr, we decided to have pictographs linked to the Unified Medical Language System (UMLS) concept rather than words. PictoDr works as an automatic translation, with a statistical and neural background. In PictoDr, the caregiver can speak as there is a vocal recognition mapping with statistical and neural analysis to the UMLS gloss (sequence of UMLS) concepts, followed by mapping of these concepts to pictographs. We observed that to match the vocal recognition to the UMLS gloss as a pivot toward the pictographs does not denaturalize the sense. The pictographs used are from ARASAAC, SantéBD (https://santebd.org/) and “homemade” ones. The UMLS gloss is linked to French variations, with a part originally from BabelDr (40). Indeed, 12,000 sentences of BabelDr are linked to UMLS and permit to feed the PictoDr database. The challenge of PictoDr is to cross cultural differences and to be understood by many cultural backgrounds, thus permitting to tailor the tool to a specific population. We also asked two doctors to say what they understood from the pictographs generated and showed a level of agreement according to Cohen’s Kappa (k = 0.661) as moderate (40). Figure 2 illustrates an example of the translation word-by-word in pictographs of the input "Etes-vous enceinte?" (“Are you pregnant?”).

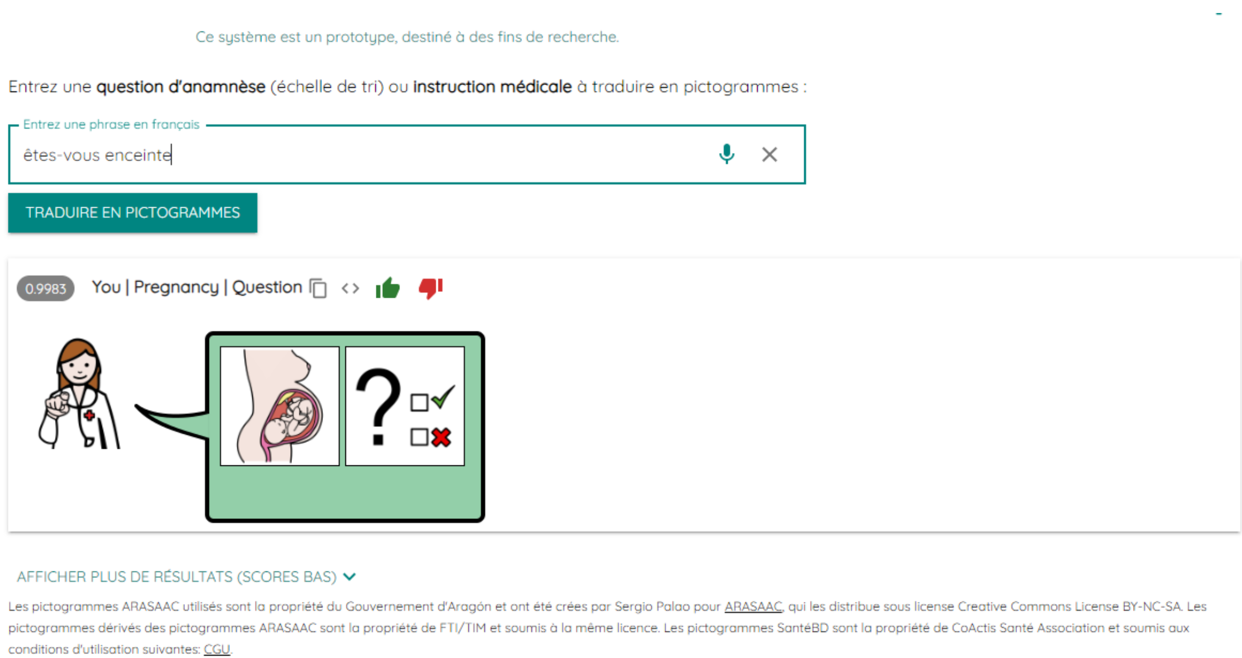


Figure 3: "Etes-vous enceinte?" ("Are you pregnant?"). (Reproduced from https://propicto.demos.unige.ch/pictodrclient/).

2. References

1. Kenny JF, Chang BC, Hemmert KC. Factors affecting emergency department crowding. Emerg Med Clin North Am. 2020;38(3):573-87.
2. Dubina MI, O'Neill JL, Feldman SR. Effect of patient satisfaction on outcomes of care. Expert Rev Pharmacoecon Outcomes Res. 2009;9(5):393-5.
3. Taylor C, Benger JR. Patient satisfaction in emergency medicine. Emerg Med J. 2004;21(5):528-32.
4. Bursch B, Beezy J, Shaw R. Emergency department satisfaction: what matters most? Ann Emerg Med. 1993;22(3):586-91.
5. Soremekun OA, Takayesu JK, Bohan SJ. Framework for analyzing wait times and other factors that impact patient satisfaction in the emergency department. J Emerg Med. 2011;41(6):686-92.
6. Thompson DA, Yarnold PR, Williams DR, Adams SL. Effects of actual waiting time, perceived waiting time, information delivery, and expressive quality on patient satisfaction in the emergency department. Ann Emerg Med. 1996;28(6):657-65.
7. Hansagi H, Carlsson B, Brismar B. The urgency of care need and patient satisfaction at a hospital emergency department. Health Care Manage Rev. 1992;17(2):71-5.
8. Grosgurin O, Cramer B, Schaller M, Sarasin FP, Rutschmann OT. Patients leaving the emergency department without being seen by a physician: a retrospective database analysis. Swiss Med Wkly. 2013;143:w13889.
9. Schwab RA. Emergency department customer satisfaction: the point of view paradox. Ann Emerg Med. 2000;35(5):499-501.
10. Sun B, Chekijian S. Emergency department crowding and other intractable problems. NEJM Catal Innov Care Deliv. 2020;1(2). doi: 10.1056/CAT.20.0061.
11. Sartini M, Carbone A, Demartini A, Giribone L, Oliva M, Spagnolo AM, et al. Overcrowding in emergency department: causes, consequences, and solutions-a narrative review. Healthcare (Basel). 2022;10(9):1625.
12. DeSai C, Janowiak K, Secheli B, Phelps E, McDonald S, Reed G, et al. Empowering patients: simplifying discharge instructions. BMJ Open Qual. 2021;10(3):e001419.

13. Showalter JW, Rafferty CM, Swallow NA, Dasilva KO, Chuang CH. Effect of standardized electronic discharge instructions on post-discharge hospital utilization. J Gen Intern Med. 2011;26(7):718-23.
14. Jacobsohn GC, Jones CMC, Green RK, Cochran AL, Caprio TV, Cushman JT, et al. Effectiveness of a care transitions intervention for older adults discharged home from the emergency department: A randomized controlled trial. Acad Emerg Med. 2022;29(1):51-63.
15. Schmutz T, Le Terrier C, Ribordy V. Flow dynamics. Scand J Trauma Resusc Emerg Med. 2024;32(1):59.
16. Nairn S, Whotton E, Marshal C, Roberts M, Swann G. The patient experience in emergency departments: a review of the literature. Accid Emerg Nurs. 2004;12(3):159-65.
17. Bjorvell H, Stieg J. Patients' perceptions of the health care received in an emergency department. Ann Emerg Med. 1991;20(7):734-8.
18. Maitra A, Chikhani C. Patient satisfaction in an urban accident and emergency department. Br J Clin Pract. 1992;46(3):182-4.
19. Spechbach H, Rochat J, Gaspoz JM, Lovis C, Ehrler F. Patients' time perception in the waiting room of an ambulatory emergency unit: a cross-sectional study. BMC Emerg Med. 2019;19(1):41.
20. Ehrler F, Lovis C, Rochat J, Schneider F, Gervaix A, Galetto-Lacour A, et al. [InfoKids: changing the patients' journey paradigm in an Emergency Department]. Rev Med Suisse. 2018;14(617):1538-42.
21. Ehrler F, Rochat J, Siebert JN, Guessous I, Lovis C, Spechbach H. Use of a semiautomatic text message system to improve satisfaction with wait time in the adult emergency department: cross-sectional survey study. JMIR Med Inform. 2022;10(9):e34488.
22. Puthenparampril E, Owens K, Huang J, Ok S, Campbell A, Ajala A, et al. Using text messaging to improve communication and patient experience in a perioperative setting. J PeriAnesth Nurs. 2023;38:e28.
23. Scolnik D, Matthews P, Caulfeild J, Williams C, Feldman BM. Pagers in a busy paediatric emergency waiting room: A randomized controlled trial. Paediatr Child Health. 2003;8(7):422-6.
24. Beggs CB, Shepherd SJ, Kerr KG. Potential for airborne transmission of infection in the waiting areas of healthcare premises: stochastic analysis using a Monte Carlo model. BMC Infect Dis. 2010;10:247.
25. Odendaal I, Tönsing KM. Augmentative and alternative communication for individuals with post-stroke aphasia: perspectives of South African speech-language pathologists. Augment Altern Commun. 2024:1-9.
26. Arimon MP, Llobet MP, Roldán-Merino J, Moreno-Arroyo C, Blanco MÁ H, Lluch-Canut T. A communicative intervention to improve the psychoemotional state of critical care patients transported by ambulance. Am J Crit Care. 2021;30(1):45-54.
27. Gierlinger S, Almulla N, Lee C, Lee T. What really matters to patients. NEJM Catalyst. 2024;5(4). doi:10.1056/CAT.24.0328.
28. van Rosse F, de Bruijne M, Suurmond J, Essink-Bot ML, Wagner C. Language barriers and patient safety risks in hospital care. A mixed methods study. Int J Nurs Stud. 2016;54:45-53.
29. Krishel S, Baraff LJ. Effect of emergency department information on patient satisfaction. Ann Emerg Med. 1993;22(3):568-72.
30. Jaeger FN, Pellaud N, Laville B, Klauser P. The migration-related language barrier and professional interpreter use in primary health care in Switzerland. BMC Health Serv Res. 2019;19(1):429.
31. Flores G, Laws MB, Mayo SJ, Zuckerman B, Abreu M, Medina L, et al. Errors in medical interpretation and their potential clinical consequences in pediatric encounters. Pediatrics. 2003;111(1):6-14.
32. Flores G, Abreu M, Barone CP, Bachur R, Lin H. Errors of medical interpretation and their potential clinical consequences: a comparison of professional versus ad hoc versus no interpreters. Ann Emerg Med. 2012;60(5):545-53.
33. Khander A, Farag S, Chen KT. Identification and evaluation of medical translator mobile applications using an Adapted APPLICATIONS Scoring System. Telemed J E Health. 2018;24(8):594-603.

34. Hudelson P, Chappuis F. Using voice-to-voice machine translation to overcome language barriers in clinical communication: an exploratory study. J Gen Intern Med. 2024;39(7):1095-102.
35. Succop B, Currin M, Hesse G, Black H, Andrews B, Elton SW, et al. Automated translation accurately translates recorded pediatric neurosurgery clinic conversations between Spanish and English. Neurosurg Rev. 2024;47(1):210.
36. Janakiram AA, Bouillon P, Gerlach J, Hudelson P, Spechbach H. [I'm having trouble communicating with my patient at the emergency unit: what tools are available?]. Rev Med Suisse. 2021;17(739):995-8.
37. Spechbach H, Gerlach J, Mazouri Karker S, Tsourakis N, Combescure C, Bouillon P. A speech-enabled fixed-phrase translator for emergency settings: crossover study. JMIR Med Inform. 2019;7(2):e13167.
38. Janakiram AA, Gerlach J, Vuadens-Lehmann A, Bouillon P, Spechbach H. User satisfaction with a speech-enabled translator in emergency settings. Stud Health Technol Inform. 2020;270:1421-2.
39. Barton-Hulsey A, Boesch MC, Chung Y, Caswell T, Sonntag AM, Quach W. Emergency and disaster preparedness for individuals who use augmentative and alternative communication: a pilot study on supported planning using a toolkit. Am J Speech Lang Pathol. 2024;33(1):16-32.
40. Bodenreider O. The Unified Medical Language System (UMLS): integrating biomedical terminology. Nucleic Acids Res. 2004;32(Database issue):D267-70.

3. Aim of the thesis

My thesis comprises two main topics represented by four articles that describe the potential strategies/instruments for improvement of the overall satisfaction in the context of fluctuating overcrowding in an outpatient emergency unit. First, I describe the analysis conducted in the outpatient emergency unit (first article described) (1) on the determinants of patient satisfaction before the medical consultation, with a focus on the waiting time before seeing a doctor. Second (second article described), I focus on how we can act on the patient's perception of this time by describing two tools developed and the implementation of the next steps, i.e., a screen in ED allowing the patient to situate him/herself in the queue, and a semi-automatic SMS alerting patients to the timing of the consultation and permitting them to wait outside the unit (2).

I will then describe the instruments (third and fourth article described) (3-4) that I have developed in collaboration with the University of Geneva Faculty of Translating and Interpreting (Professor Pierrette Bouillon), sustained by the private foundation of Geneva in order to have a reliable instrument (www.babeldr.unige.ch) permitting to communicate with people with whom we do not share a common language. I will also discuss the use of French sign language (in French-speaking Switzerland) and the subsequent extension of BabelDr, named PictoDr (https://propicto.demos.unige.ch/pictodrclient/). The latter has the great advantage to allow a two-way communication (patient and caregiver) with pictographs, thus freeing us from the need to use language and permitting an augmentative communication (5). Regarding communication during the medical consultation with these tools, I will focus on the satisfaction of both the patient and caregiver.

3.1 Own contributions in the area of research through the presentation of 4 articles

3.2 What are the instruments developed to optimize the wait perception before seeing a doctor and the level of satisfaction (based on two studies conducted in the HUG outpatient emergency unit? (1-2)

As previously discussed, if the wait perception is improved as in the model “satisfaction = perception – expectation”, this should improve in turn the overall patient satisfaction with care in the ED. We propose a solution impacting on the wait perception, which offers two wait possibilities, i.e., the patient either stays in the waiting room or outside in another location.

3.2.1 First article: "Patients’ time perception in the waiting room of an ambulatory emergency unit: a cross-sectional study" (1)

We focused on the factors influencing the perception of time until medical contact through a survey with the hypothesis that this will impact on satisfaction, but without including other parameters of quality, such as an appropriate diagnosis.

Five hundred and nine patients with a wide variety of reasons for visiting the ED completed the survey in the ED waiting room. Most (88.5%) were classified with an emergency level of 3. The inclusion criterion was being capable of completing the survey in French. The survey had 27 items evaluated using a 5-point Likert scale (“very long”, “long”, “acceptable”, “short” and “very short” for the time) and some items with time in minutes (such as "perceived" or "expected" waiting time). Actual waiting time (the reality) was collected through our informatics system. We focused on the most frequent factors influencing waiting time perception and satisfaction in the ED as identified in the literature (6).

We found that the actual waiting time significantly influences the perception of wait. The wait perception was evaluated as “very short”, “short” and “acceptable” under an exact waiting time mean of 1.03h (standard deviation = 0.5). The actual waiting time discriminate value was p <0.001 between the five categories of perceived time (“very short”, “short”, “acceptable”, “long”, “very long”). If the delta between the actual waiting time and the expected waiting time (completed by respondents upon arrival in the ED) were >1, the perceived waiting time was estimated as “very long” or “long” and had a direct negative impact on the overall satisfaction (negative disconfirmation paradigm) (7).

In an ordered regression for the outcome variable "perceived waiting time as very short", the most influencing factor was correct triage as already reported (8) (OR] 5.37 [95% CI, 2.46-11.73]; p<0.001). The second factor was the “feeling of being forgotten” (OR, 3 [95% CI, 1.47-6.15]; p<0.001) (the question was “I didn’t feel forgotten”), followed by non-respect of privacy (OR, 0.21 [95% CI, 0.05-

0.88]; p=0.03), and the actual waiting time (OR, 0.98 [95% CI, 0.98-0.99]; p=0.03). After a 30-minute wait, patient's "perceived waiting time as very short" decreased by 81% and every additional 1 minute of actual waiting time was associated with a 2% decrease in the chance that patients would perceive the waiting time as very short. These items are consistent with the published literature (6).

Patients' expected mean of an acceptable waiting time before seeing a doctor was 65.42 minutes as already reported in the literature by Shaikh et al (9), who also found that people want to leave before seeing a doctor after 2 hours of wait time. It is also known that overall satisfaction is strongly associated with the perception of the wait to see the doctor being shorter than expected (10).

Notably, we used this analysis to develop new plans to improve the architecture and organization of our unit. Following the results of this study, we placed a screen in the waiting room allowing patients to identify themselves in the queue with "bubbles" containing their initials (Figure 4). We have also conducted a post-display study to measure the same items as before installation of the screen. Data collection is currently ongoing.

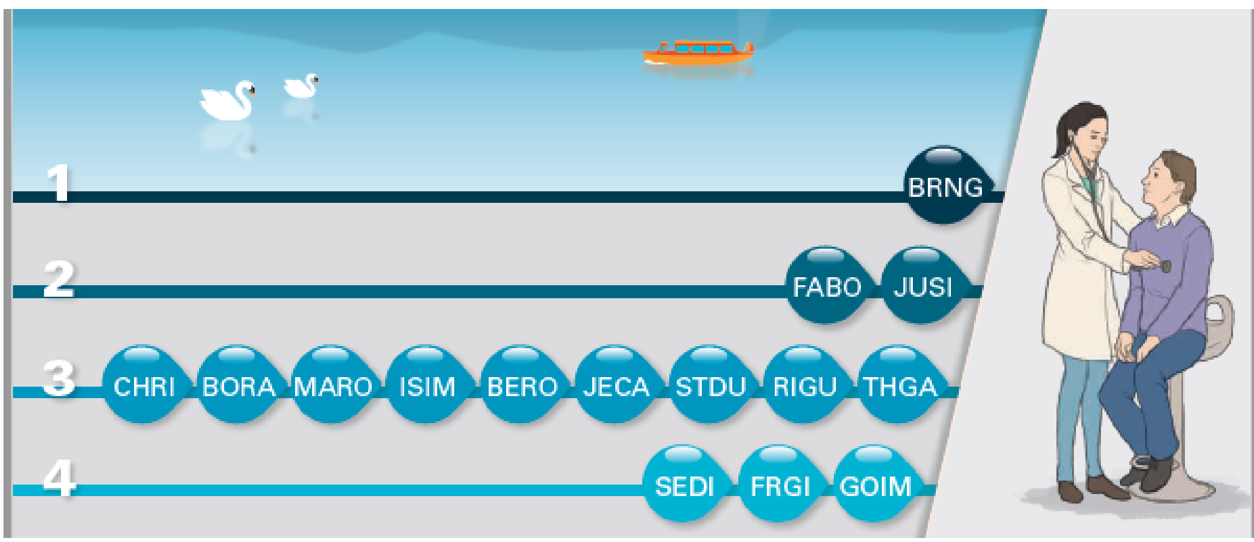


Figure 4: Bubbles containing patients' initials. (Reproduced from the application Infomed in the DPI (Dossier Patient Informatique [electronic patient record]) at HUG.)

Spechbach et al. BMC Emergency Medicine (2019) 19:41
https://doi.org/10.1186/s12873-019-0254-1

BMC Emergency Medicine

RESEARCH ARTICLE Open Access

Patients' time perception in the waiting room of an ambulatory emergency unit: a cross-sectional study

Hervé Spechbach[1], Jessica Rochat[2,3], Jean-Michel Gaspoz[1], Christian Lovis[2,3] and Frederic Ehrler[3*]

Abstract

Background: Patient satisfaction has become an increasingly important element in a service-oriented healthcare market. Although satisfaction is influenced by many factors, the waiting time to be seen by medical staff has been shown to be one of the key criteria. However, waiting is not an objective experience and several factors can influence its perception.

Methods: We conducted a questionnaire-based, cross-sectional study among patients attending the emergency unit of a Swiss university hospital in order to explore the key factors influencing wait perception.

Results: A total of 509 patients participated in the study. Appropriate assessment of emergency level by caregivers, the feeling of being forgotten, respect of privacy, and lack of information on the exact waiting time were identified as significant variables for wait perception.

Conclusions: Our study confirmed the existence of a 'golden hour' when the patient is willing to wait until the medical encounter. In case the wait cannot be limited, an appropriate assessment of the emergency level by caregivers and avoiding the patients of feeling being forgotten are very important factors to avoid a negative perception of the waiting time before seeing a doctor.

Trial registration: (ID REQ-2016-00555).

Keywords: Patient satisfaction, Emergency department, Waiting time, Perceived waiting time, Service-oriented health care

Background

The emergency department (ED) of a health facility provides services whose evaluation is strongly linked to the 'customer' experience [1]. In a service-oriented healthcare market, patient satisfaction is a quality indicator receiving increasing attention. Dissatisfied patients are likely to share their negative experience with other people, thus having a negative influence on service perception [2, 3]. In addition, internet promotes a rapid and wide dissemination of these opinions. This word-of-mouth marketing is powerful, especially as consumers grow more knowledgeable about their healthcare choices. More importantly, there is evidence of a reciprocal relationship between patient satisfaction and continuity of care, which is associated with improved patient outcomes. For instance, satisfaction with care has been shown to have a significant influence on patient adherence with treatment plans [4].

It is difficult to explore the holistic experience of being a patient as consideration must be given to the multiple facets of his/her psychological behaviour related to the cultural aspect of care [5]. Waiting time is considered to be an important determinant of patient satisfaction and should be carefully considered. Bursch et al. [6] showed that the duration of waiting time before ED care, patients' evaluation of care by doctors and nurses, the organization of ED staff, and quality of received information were the most important variables associated with an overall satisfaction with ED services. A Swedish study [7] including 758 patients investigated patient satisfaction with treatment and service at an ED. Results showed that satisfaction was lower among

* Correspondence: frederic.ehrler@hcuge.ch
[3]Division of Medical Information Sciences, Geneva University Hospitals, Geneva, Switzerland
Full list of author information is available at the end of the article

patients who were triaged as non-urgent (and thus waited longer) than among the immediate and urgent triaged patients [8]. In the USA, the Kaiser Permanente's integrated health care model [9] identified that the most important variable associated with overall satisfaction with ED services was satisfaction with the amount of time spent waiting before the patient was cared for in the ED.

If the actual waiting time is a strong factor for satisfaction, its perception is also a key factor. Marketing and consumer research literature [10] found that satisfaction is mainly influenced by the difference between the consumers' acceptable and perceived waiting time, with a linear relation between actual and perceived waiting time [11, 12]. Similarly, several international studies conducted in the ED showed that satisfaction is strongly correlated with perceived waiting time and its correlation with the actual waiting time. Waiting time at the ED is often perceived as unreasonably long and information received from nurses and doctors is judged to be unsatisfactory [13, 14]. Among the factors influencing the perceived waiting time, the role of the expected waiting time can be highlighted. The perceived waiting time influences the satisfaction we have from a service through our expectation (i.e. the expected waiting time) and is explained by the disconfirmation paradigm perceptions of a service [15] in that dissatisfaction arises when service expectations are not met (i.e. a positive disconfirmation occurs when a service is perceived as being better than expected and a negative disconfirmation occurs when a service is perceived as being worse than expected). Growing pressure to provide care for more patients over the last decade has led to overcrowding and longer waiting times [16], thus leading to dissatisfaction from ED patients. Moreover, it has been demonstrated that the longer a patient waits before receiving care, the more s/he is at risk of leaving without being seen [17]. In this context, it is critical to understand clearly the factors influencing the perception of wait of ED patients.

In Switzerland, a consultation at the ED of a public hospital results in a charge of several hundred francs to patients who pay one part, the other part being reimbursed by their health insurance. Therefore, patients expect a certain level of service and may complain when dissatisfied (0.3% of visits). Among patient complaints, those concerning the waiting time until a medical contact are recurrent at Geneva University Hospitals, although not frequent (0.1% of visits), similar to other centres elsewhere [18–20]. As a consequence, the hospital management has committed itself to improving the situation and thus enhance the institutional image among the local population. In this context, our study aims to investigate the different factors influencing the perception of waiting time until medical contact through a questionnaire distributed to patients triaged as non-urgent in the outpatient emergency unit. We did not include other parameters of quality of care in our analysis, such as the actual medical care or appropriate diagnosis and follow-up, and focused only on the emergency unit triage and the initial patient throughput until a medical contact.

Methods

Study design

We conducted a questionnaire-based, cross-sectional, descriptive study between 17 December 2015 and 14 March 2016.

Setting

The survey was conducted at the outpatient emergency unit of Geneva University Hospitals, the largest university hospital in Switzerland with over 1900 beds. The unit is separated into 'orange' and 'green' subunits. The 'orange' subunit cares for general ambulatory pathologies and the 'green' subunit for surgical ambulatory pathologies. Both have six consultation rooms each. Each subunit has its own semi-closed waiting room with seating, a television, water and newspapers. The staff (doctors and nurses) are the same for the entire unit and work part of the week in the 'orange' subunit and another part in the 'green' subunit. Both units open at 8 h00 and close at 23 h00. After 21 h00, if more than six patients are already being cared for and a further six are in the waiting room, the units are considered as closed and all patients arriving after that time remain in the inpatient emergency department, which is open 24 h/24 h. At 23 h00 all patients present in the outpatient emergency unit are transferred to the inpatient unit. In 2016, 22,000 patients consulted the unit (half in the 'orange' and half in the 'green'). Median length of stay is 3 h in the 'orange' subunit and 2.8 h in the 'green' subunit.

When a patient arrives at the emergency entrance, s/he is first seen by a triage nurse who decides if the patient is a candidate for the ambulatory emergency unit, based on the Swiss Emergency Triage Scale (SETS). Level 1 is a life −/limb-threatening situation where the patient must be seen by a medical doctor immediately, level 2 in the following 20 min, level 3 in 120 min, and level 4 is considered as non-urgent. Eighty percent of patients visiting the emergency unit are classified as level 3 and 10% as level 4. After triage, the patient goes through an administrative registration process and is then directed to one of the subunits by following coloured lines on the floor. These lead to a nursing desk where a nurse accompanies the patient to the waiting area. Both subunits have their own semi-closed waiting room. Whenever possible, nurses inform patients about the waiting time using their own estimation. As soon as a consultation room and a doctor are available, the patient is taken to the room by the nurse. After the medical consultation, the patient can either return home or may have to undergo an additional examination and return again to the waiting room. A small percentage of patients (5%) are hospitalized and 5% leave the unit without being seen by a doctor.

Population

Patients presenting to the outpatient emergency unit were invited to participate to the study if they were at least 16 years old and French-speaking. Exclusion criteria were patients not capable of discernment (e.g. unconscious, under the influence of drugs, suffering extreme trauma or with cognitive disorders), unable to read/understand French, vision problems, those who had already completed the questionnaire, and patients in severe pain or too aggressive. The patient journey until data collection is shown in Fig. 1.

Questionnaire

The questionnaire developed for this study was based on the most frequent factors related to waiting time perception and overall satisfaction in the ED identified in the literature [21–29] regarding expected waiting time: the patient's estimate of the maximum reasonable waiting time before seeing a doctor; perceived waiting time; appropriate assessment of emergency level by caregivers; perception of their own emergency level compared to other patients; satisfaction with information received; organization of the emergency unit; feeling of being forgotten; frustration; capability to position themselves in the waiting queue; respect of privacy; comfort in the waiting room; and an indication of waiting time by the nurse. Additional items were added to explore specific topics of interest, such as the recommendation of the emergency unit to others, the interest in waiting elsewhere and overall satisfaction. The final questionnaire consisted of 27 items. Most items were evaluated with a 5-point Likert scale ranging from 'strongly disagree' to 'strongly agree'. The item concerning the perceived waiting time was evaluated with a 5-point Likert scale ranging from 'very short' to 'very long'. The item concerning the perception of their own emergency level compared to other patients was evaluated with a 3-point Likert scale ranging from 'superior' to 'inferior'. The two items concerning the expected waiting time and the patient's estimate of the maximum waiting time before seeing a doctor were questions requiring a time input in hours and minutes. To collect the real waiting time, we extracted the time that the patient actually saw the doctor from the clinical information system (CIS). Indeed, each time a doctor starts the medical care, s/he notifies the CIS by clicking a specific button and thus generates a log into the system

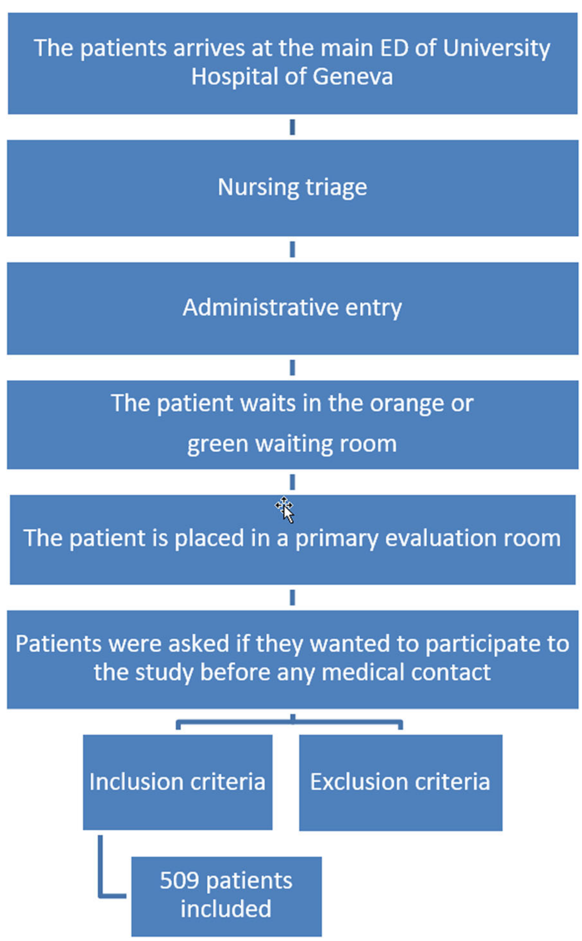


Fig. 1 Study flow chart

Procedure and ethical considerations

The institutional ethics committee approved the study protocol. Participation to the study was voluntary and by oral consent. Questionnaires were distributed over a period of three months by a total of 20 nurses during two different shifts: 7 h30 to 16 h00 and 15 h00 to 23 h30.

Following a convenience sampling approach, each questionnaire was distributed to a patient by a nurse in a consultation room while waiting for the doctor. We did not include patients who did not wait before going to a consultation room. Following verification of the inclusion criteria, the nurse asked the patient if s/he agreed to complete the questionnaire. Information about the study and confidentiality were given orally. The survey took approximately 20 min to complete. The nurse remained available for any questions and to help the patient complete the questionnaire if necessary. Once completed, questionnaires were collected by the nurses and placed in a dedicated box in a secure room. Questionnaires were collected by a scientific collaborator each morning and entered into an Excel file. In order to link the questionnaire data to the data extracted from the CIS, we used a mapping file linking the ID of the questionnaire to the patient ID. Once all data were included in the Excel file, only the questionnaire ID was retained to ensure anonymous analysis.

Statistical analysis

Data were analyzed using Stata/IC 14 software [30]. Descriptive statistics and frequencies were produced to describe the demographic and medical characteristics of participants. Variables evaluated with a 5-point Likert scale were re-coded into three categories: 'disagree,' 'neither disagree nor agree,' and 'agree'. In addition, the actual waiting time of each patient that responded to the questionnaire was extracted from the hospital CIS. The variable 'difference between actual and expected time' was created by subtracting the duration that the patient thought that s/he should have waited before being seen by a doctor from the actual duration of the patient wait. If the value was superior to zero, it meant that the patient waited longer than expected. Statistically significant predictors of the waiting time perception were determined by an ordered logistic regression using every variable as input. Missing data were removed from the regression analysis.

Results

A total of 509 patients participated in the study. The demographic characteristics of participants are presented in Table 1.

Table 2 shows the medical characteristics of respondents. Most respondents (88.5%) were classified with an emergency level of 3, 3.5% were level 2, and 8% were triaged as level 4 (the less urgent).

Relation between wait perception and exact waiting time

Figure 2 illustrates the wait perception according to the actual waiting time and suggests that the perception of waiting time is directly correlated to its duration. Patients perceiving the wait as very long were those who waited the most on average. A non-parametric ANOVA (Kruskal-Wallis test) was performed to verify if the mean differences of the actual waiting time according to the five modalities of the wait perception (i.e. very short, short, acceptable, long, very long) were statistically significant. The non-parametric test showed a statistically significant difference (Chi-square = 220; $p < 0.001$). To define which differences were significant, 10 multiple comparisons were performed using the Mann-Whitney test, taking into account the Bonferroni correction. All differences were statistically significant with a p-value inferior to 0.001. This result proves that the actual waiting time influences significantly the perception of the wait. Figure 2 shows that wait perception is considered as very short, short and acceptable under an exact waiting time mean of 1.03 h (standard deviation = 0.50).

Table 1 Demographic characteristics of respondents

Demographic characteristics	Respondents
	Mean (standard deviation)
Age (years)	42 (17)
	n (%)
Sex	
Male	266 (52.3)
Female	243 (47.7)
Nationality	
Swiss	268 (52.7)
French	63 (12.4)
Portuguese	44 (8.6)
Spanish	17 (3.3)
Italian	15 (2.9)
Others	102 (20.1)

Table 2 Medical characteristics of respondents

Medical characteristics	Respondents
	n (%)
Emergency level according to SETS	
1	0 (0)
2	18 (3.5)
3	450 (88.4)
4	41 (8.1)
Reason for ED visit	
Limb trauma	128 (25.1)
Skin/soft tissue ailment/infection	54 (10.6)
Arthralgia, myalgia, neuralgia	45 (8.8)
Pain and/or oedema of a limb	36 (7.1)
Abdominal pain	24 (4.7)
Cervicalgia, back pain, low back pain	24 (4.7)
Influenza syndrome	24 (4.7)
Superficial wounds	17 (3.3)
Cough, sputum	15 (2.9)
Others	142 (28.1)
Family doctor	
Yes	379 (74.5)
No	127 (25)
Missing data	3 (0.5)
Already consulted at the same emergency unit	
Yes	355 (69.7)
No	154 (30.3)

SETS Swiss Emergency Triage Scale

Relation between wait perception and wait confirmation

The difference between the actual and expected waiting time, known as 'wait confirmation,' measures the difference between the expectation and the reality. Patients whose actual wait exceeded the expected waiting time perceived the wait as very long (Fig. 3). A non-parametric ANOVA (Kruskal-Wallis test) was performed to verify if the mean differences between the delta actual-expected waiting time

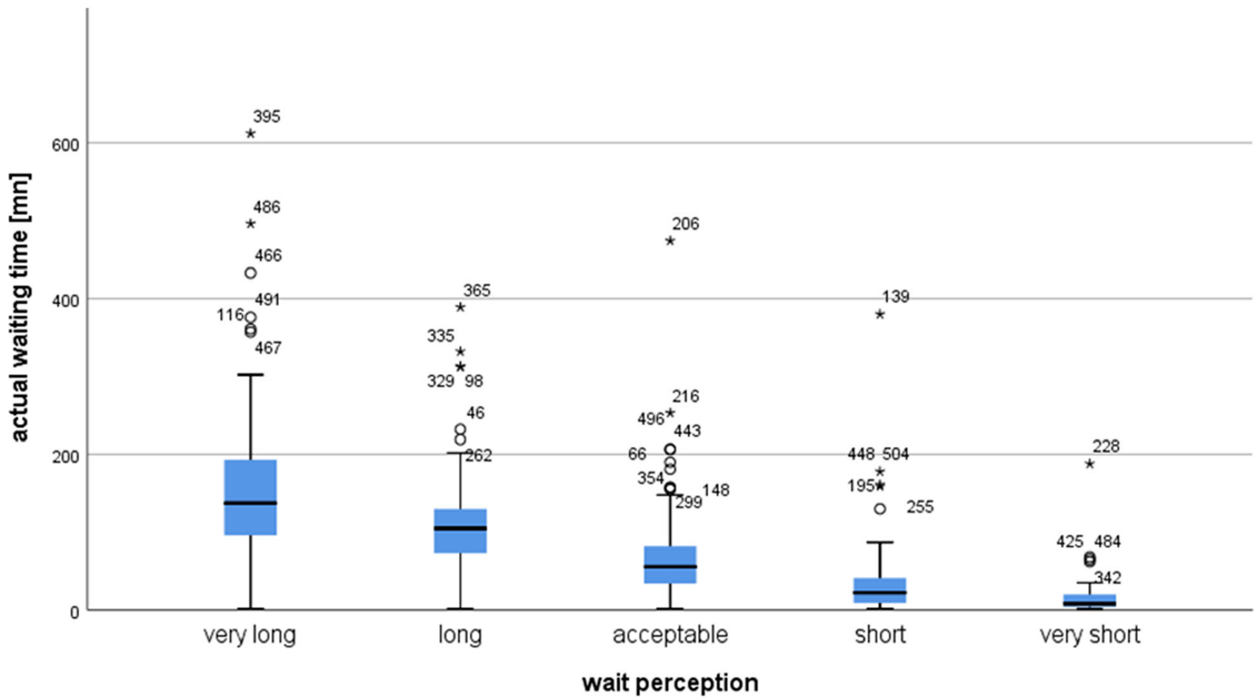


Fig. 2 Average waiting time according to wait perception

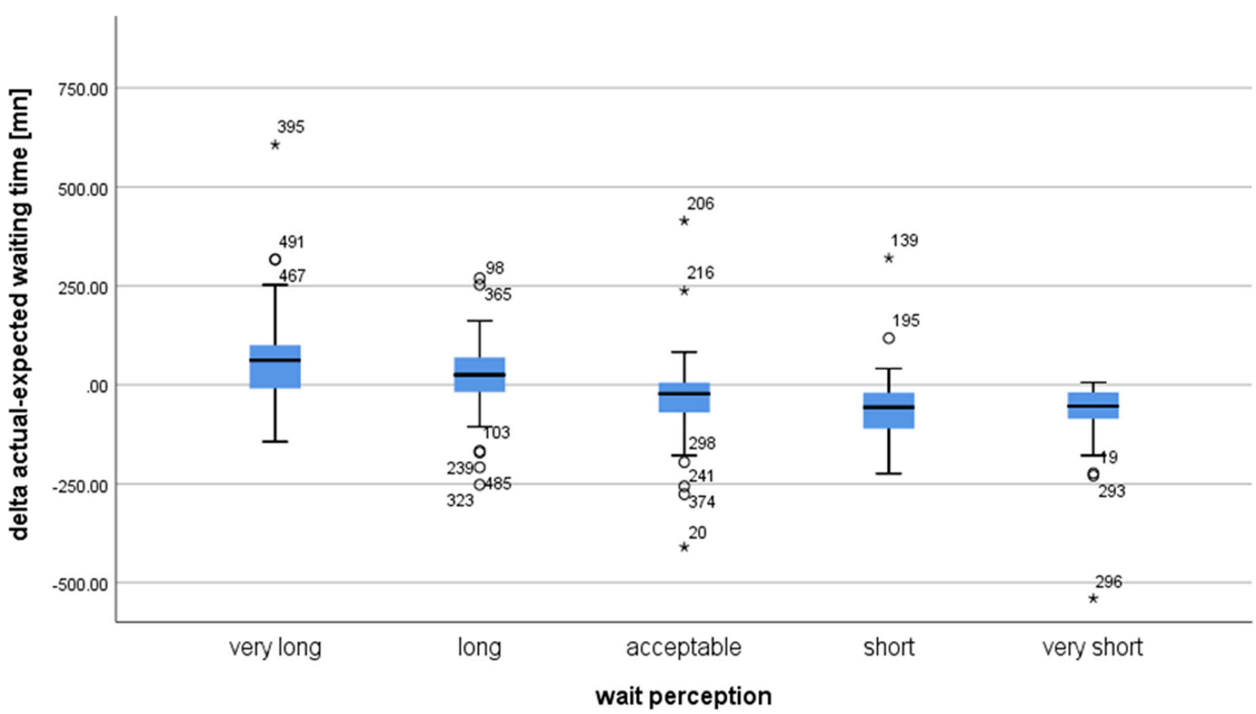


Fig. 3 Average difference between the actual and expected waiting time according to the wait perception

according to the five modalities of the wait perception (very short, short, acceptable, long, very long) were statistically significant. The non-parametric test showed a statistically significant difference (Chi-square = 112; $p < 0.001$). To define which differences were significant, 10 multiple comparisons were performed using the Mann-Whitney test, taking into account the Bonferroni correction. Apart from the difference between the 'very long' and 'long' categories, all differences were statistically significant with a *p*-value inferior to 0.001. This result proves that the difference between the actual and expected time significantly influences the wait perception.

Ordered logistic regression of the wait perception

We performed an ordered logistic regression to identify the factors influencing the wait perception (Table 3) on the 325 observations without missing data. All variables that were statistically significant in univariate analysis were integrated into the multivariate model. The reference value was 'neither disagree nor agree', except for the variable 'exact waiting time', which was continuous.

The outcome variable was the perception of the waiting time and the reference category was 'very short'. The most influencing factor of the wait perception was the appropriate assessment of the patient emergency level by caregivers. The odds of perceiving the wait as 'very short' were 5.37 times (95% CI 2.46–11.73; $p < 0.001$) greater for a patient considering their emergency level adequately assessed compared to a patient answering 'neither disagree nor agree'. The second factor was the 'feeling of being forgotten'. In Tables 4, 23% (115/499) of respondents reported having felt forgotten. The odds of perceiving the wait as 'very short' were three times (95% CI 1.47–6.15; $p < 0.001$) greater for a patient that did not feel forgotten in the waiting room compared to a patient answering 'neither disagree nor agree'. The other predictor was 'respect of privacy'. Table 4 shows that 5.8% (29/495) reported that their privacy was not respected in the waiting room. The odds of perceiving the wait as 'very short' were 4.76 times lower (odds ratio [OR] = 0.21; 95% CI 0.05–0.88; $p = 0.03$) for a patient considering that his/her privacy was not respected compared to a patient answering 'neither disagree nor agree'.

The last predictor was the actual waiting time with an OR of 0.98 (95% CI 0.98–0.99; $p < 0.001$). Every one minute increase in the actual waiting time is associated with 2% decrease in the chance of the patient perceiving the wait as 'very short'. In other words, after 30 min of waiting, patients' perceived waiting time as 'very short' would decline by 81%.

Discussion

Our results showed that 41% of respondents expressed agreement on satisfaction in the waiting room and 30% strongly agreed. Mean waiting time expected by the patient was 94.85 min. Patients' estimation of the mean reasonable waiting time before seeing a doctor was 65.42 min, with a mean of 78 min for the actual waiting time. Twenty-four percent of respondents described their perception of the waiting time as being short, 43.5% as acceptable, and 32.4% as long. Indeed, the average waiting time was considered as very short, short and acceptable until a mean of 63 min. Interestingly, 53.2% reported that they did not receive information from caregivers about the waiting time. The model used for this study (annex 1) identified four significant predictors of the wait perception in decreasing order: appropriate assessment of emergency level by caregivers; feeling of being forgotten; respect of privacy; and the exact waiting time. Of note, 23% (115/499) of respondents reported to have felt forgotten. Regarding the appropriate assessment of emergency level by caregivers, only 4.5% said that they disagreed ('strongly disagree' and 'disagree'). Similarly, 8.2% disagreed with the organization of the emergency unit; 67.5% were interested to wait elsewhere, and 84.3% said that their privacy was respected ('agree' and 'strongly agree').

Appropriate assessment of emergency level by caregivers

Accurate assessment of the emergency level by caregivers is the strongest influencer of the wait perception. Reasons for attending EDs are linked, but not only to the perception of situation urgency. Some patients in Nederland, Australia also mentioned that they went to the ED to be able to see a doctor and have any tests or X-rays all done in the same place or because they did not want to wait for an appointment with their general practitioner [31, 32]. Furthermore, nurses and doctors may disagree about the patient's triage category [33]. The discrepancy in severity assessment between caregivers and patients has also been highlighted in the research of Toloo et al. [17]. In their study, they reported that almost 50% of patients had the feeling of being undertriaged and 20% expected a lower priority than the actual triage category. The correlation between the perceived priority and actual triage category was weak. In a patient-centred approach, caregivers should be encouraged to explain clearly to patients the reasons for their classification at a given level and why other patients may be possibly seen before them by medical staff. Therefore, it would be important to create different support materials that allow the patient to understand the reasons for his/her assigned triage category. In addition, more consideration should perhaps be given to the perceived urgency by patients as Toloo et al. [17] showed that this is associated with an expected higher priority triage.

Feeling of having been forgotten

The feeling of having been forgotten is identified as the second strongest influencing factor of the wait perception. Although this criteria was not retrieved in our literature

Table 3 Ordered logistic regression of the wait perception: the significant factors

	Univariate analysis				Multivariate analysis			
	Odds ratio	$p > \|z\|$	[95% CI]		Odds ratio	$p > \|z\|$	[95% CI]	
Perception of the waiting time								
Appropriate assessment of emergency level by caregivers								
Agree	12.07	< 0.001	5.15	28.30	5.37	< 0.001	2.46	11.73
Feeling of being forgotten								
Disagree	5.2	< 0.001	3.18	8.49	3.00	< 0.001	1.47	6.15
Respect of privacy								
Disagree	0.26	< 0.001	.11	.59	.21	0.03	0.05	0.88
Exact waiting time	0.98	< 0.001	.97	.98	.98	< 0.001	0.98	0.99

Table 4 Descriptive statistics

Variable	Strongly disagree (%)	Disagree (%)	Neither agree nor disagree (%)	Agree (%)	Strongly agree (%)	Missing data (N)
Satisfaction in the waiting room	4.8	6.2	17.9	41.0	30.0	12
Waiting time matches patients' expectations	12.6	20.3	13.2	35.4	18.5	17
Appropriate assessment of emergency level by caregivers	0.8	3.7	15.4	51.0	29.1	17
Satisfaction with information received about the stages of care in the emergency unit	1.6	3.0	12.7	52.2	30.5	11
Organization of the emergency unit	2.2	6.0	16.0	52.5	23.4	8
Feeling of being forgotten	31.5	30.1	15.4	16.0	7.0	10
Frustration caused by patients being seen before you	31.6	28.1	25.9	9.1	5.2	104
Being able to position yourself in the waiting queue	10.1	12.5	29.5	34.1	13.8	52
Interested in waiting elsewhere	7.9	8.3	16.3	31.1	36.4	17
Respect of privacy	2.4	3.4	9.9	47.9	36.4	14
Discomfort disrupted by other patients	39.8	28.3	17.5	10.7	3.7	22
Discomfort disrupted by the coming and going of caregivers	45.6	30.0	16.0	5.3	3.1	22
Discomfort related to the material/furniture	38.1	28	17.5	11.3	5.1	23
Discomfort disrupted by the lack of occupations/distractions	28.5	26.8	23.5	15.7	5.6	24
Would recommend the emergency unit to others	2.0	5.4	15.9	41.3	35.3	13
	Superior (%)	Similar (%)	Inferior (%)	Missing data		
Patient perception of their own emergency level compared to other patients	9.2	68.7	22.1	42		
	Not informed (%)	Informed	Cannot remember (%)	Missing data (N)		
Indication of a waiting time by the nurse	53.2	35.8	11.0	18		
	Mean	Standard deviation	N	Missing data (N)		
Waiting time expected by the patient (min)	94.85	63.68	481	28		
Patient's estimation of the maximum reasonable waiting time before seeing a doctor (min)	65.42	47.97	472	37		
Actual waiting time of the patient (minutes)	78	77	508	1		
	Very long (%)	Long (%)	Acceptable (%)	Short (%)	Very short (%)	Missing data (N)
Perception of the waiting time	13.1	19.3	43.5	14.1	9.9	6

review on the topic, we found one article in a related domain by Gilmartin et al. [34] who reported that patients felt abandoned during preoperative wait in one care centre. In this study, the main reasons of feeling abandoned was the lack of information about the delay, the process, and poor interaction with caregivers. In our setting, when a patient is seated in the waiting room, s/he usually has to wait until a nurse comes to take him/her to an examination room. During this entire time, the patient is isolated and often lacks information and is therefore unsure whether s/he has been forgotten by staff.

Link between the wait perception and actual time

The statistical analysis showed that the actual waiting time influences significantly its perception and that patients are ready to wait up to 1 h on average before considering the wait as excessive. This finding is supported by Sanober et al. [35] showing that patients were willing to wait up to 2 h before leaving the ED without being seen. Another study [36] demonstrated that patients felt that they should be seen within 1 h on average, but expected to wait 2.1 h. After 2 h, people wanted to leave the ED before seeing a medical doctor.

Eighty percent of patients visiting our emergency unit were classified as level 3 and must be seen by a doctor within a delay of 120 min according to the Swiss Emergency Triage Scale. Our observations highlighted that we must either reduce this waiting time or act on it by changing the wait perception. For example, it is known that the perception of waiting time, efficiency and the clinical skills of the emergency doctor is improved with periodic personal interaction and the provision of clinically-based information [37].

Link between the confirmation and perceived time

As described by Thompson et al. [9, 13], we showed that the wait perception was correlated with the discrepancy between the reality and the expectation of the wait. A gap between performance and expectations generates a 'disconfirmation'. Moreover, it is known that if a patient expects to wait longer than the actual wait, s/he is more satisfied, independent of the length of the waiting time [5]. Antinides et al. [10] showed that waiting time 'fillers', such as repeated information about wait duration, length of the queue, and music positively influence the time perception. This should also encourage caregivers to provide an overestimation of waiting time. This mechanism has been observed in other areas, such as waiting for transportation, where providing real-time information has demonstrated a positive influence on time perception. In our emergency unit, nurses give a personal estimate of the waiting time to at least 36% of patients.

Respect of privacy

Respect of privacy was the least significant predictor. The patients in our units wait in a semi-closed room with a television, bottles of water and a choice of different newspapers. Nurses come regularly to the waiting room to check the pain scale and sometimes provide information on an estimate of waiting time. The design of the waiting environment can be a significant factor in improving patient satisfaction. In a questionnaire on patients' privacy and satisfaction in an urban university-based hospital ED [38], 75% agreed and strongly agreed that privacy was very important for their emergency care. Following these results, an intervention including redesigning the ED environment, process management and staff education was implemented and showed significant improvements with patient perceptions on increased privacy and satisfaction. In our study, 84% of patients agreed and strongly agreed that their privacy was respected. Questions regarding discomfort showed that most patients (> 60%) strongly disagreed/disagreed and considered that their comfort was disrupted by other patients, by the coming and going of caregivers, or by the lack of distractions.

Study limitations

Our study has some limitations. The questionnaire was developed specifically for this study as no existing questionnaire in the literature allowed us to evaluate the perception of waiting time. As the questionnaire validity and reliability have not been tested, our results cannot be generalized without additional evaluation. The questionnaire was only available in French and excluded patients included those who were in too much pain, unable to read French, with ocular problems or who did not want to complete it. It is also possible that the nurses did not give the questionnaire during busy shifts. In addition, as the questionnaires were distributed by nurses, the social desirability effect could have had an impact on the patient's decision to participate and on their responses [39]. Another limitation is that the analysis was only in one unit of our ED. No response rate was calculated. Due to financial constraints, the questionnaire could only be distributed during a three-month period and this did not allow to reach the 700 participants targeted with the sample size calculation based on the rule of event per variable of 50. However, we did reach the minimal recommended sample size of 500 participants when using logistic regression for observational studies [40]. Of note, our evaluation took place at the end of the waiting process and did not take into account the quality of care dispensed. As this can influence overall patient satisfaction at the end of the entire process, it can be considered as a limitation.

Conclusion

In emergency care, medical consultations are not scheduled and the patient must almost always wait, especially non-urgent cases. Overcrowding of EDs is a major problem worldwide, with a negative impact on the quality of care. In the throughput process, the waiting time until the medical contact impacts on the overall patient satisfaction. Our study revealed that during the wait until the medical encounter, there is a 'golden hour' when the patient is willing to wait and a perceived waiting time of under 60 min is acceptable. This is consistent with the results of studies showing that patients may leave the ED after 2 h without seeing a doctor. Therefore, it can be recommended to provide information to patients if they have been waiting for more than 1 h.

The two strongest predictors of the perceived waiting time before seeing a doctor were appropriate assessment of the emergency level by caregivers and the feeling of being forgotten. Thus, ED staff should be particularly attentive to act on these factors. A solution could be to perform a more patient-centred triage, with more basic information and different support materials to allow the patients to understand the process according to their sociocultural background. Further research is needed to understand the psychological profile of patients who over- or underrate their urgency, as well as a more detailed analysis to discover the needs of patients who are waiting before the first medical contact.

Some of our results revealed avenues for improvement. For example, 47.9% of patients were unable to situate themselves in the waiting queue and 67.5% reported an interest to wait outside the waiting room (cafeteria, at home, somewhere else). These facts informed us that we should focus on interventions to improve the perceived waiting time and reduce the feeling of being forgotten. Consideration could be given also to moving the wait: the patient could wait elsewhere or be seen rapidly by a doctor and then wait (move the wait after the first medical contact). Furthermore, it could be interesting to retest the wait perception at the time of patient discharge as it would permit a comparison with the waiting perception at the beginning of the process before seeing a doctor and help to monitor the impact of any interventions.

Abbreviations

CIS: Clinical information system; ED: Emergency department; SET: Swiss Emergency Triage Scale

Acknowledgements

We thank Michele Le Bras and Etienne Satin for motivating the nurses to distribute the questionnaire, and Rosemary Sudan for editorial assistance.

Authors' contributions

HS, JR, J-MG, CL and FE conceived and designed the study. HS, FE and JR participated in the analyses and interpretation of results and writing of the manuscript. HS and FE were responsible for the overall conduct and organization of the study. All authors critically revised the script, approved the final manuscript, and agreed to be held accountable for all aspects of the work.

Funding

The study received no funding.

Availability of data and materials

Supporting data can be accessed at the Springer Nature repository.

Ethics approval

The study has received an exemption from the Regional Research Ethics Committee for Geneva University Hospitals (ID REQ-2016-00555). Each patient provided oral consent to participate in the study.

Consent for publication

Not Applicable.

Competing interests

The authors declare that they have no competing interests.

Author details

[1]Division of Primary Care Medicine, Department of Community Medicine, Primary Care and Emergency Medicine, Geneva University Hospitals, Geneva, Switzerland. [2]Faculty of Medicine, University of Geneva, Geneva, Switzerland. [3]Division of Medical Information Sciences, Geneva University Hospitals, Geneva, Switzerland.

Received: 3 November 2018 Accepted: 19 July 2019
Published online: 01 August 2019

References

1. Eitel DR, Rudkin SE, Malvehy MA, Killeen JP, Pines JM. Improving service quality by understanding emergency department flow: a white paper and position statement prepared for the American Academy of emergency medicine. J Emerg Med. 2010;38:70–9 PMID: 18514465.
2. Nyer PU, Gopinath M. Effects of complaining versus negative word of mouth on subsequent changes in satisfaction: the role of public commitment. Psychol Mark. 2005;22:937–53.
3. East R, Uncles MD, Romaniuk J, Lomax W. Measuring the impact of positive and negative word of mouth: a reappraisal. Australas Mark J. 2016;24:54–8.
4. Dubina MI, O'Neill JL, Feldman SR. Effect of patient satisfaction on outcomes of care. Expert Rev Pharmacoeconomics Outcomes Res. 2009;9:393–5 PMID: 19817520.
5. Nairn S, Whotton E, Marshal C, Roberts M, Swann G. The patient experience in emergency departments: a review of the literature. Accid Emerg Nurs. 2004;12:159–65 PMID: 15234713.
6. Bursch B, Beezy J, Shaw R. Emergency department satisfaction: what matters most? Ann Emerg Med. 1993;22:586–91 PMID: 8442550.
7. Hansagi H, Carlsson B, Brismar B. The urgency of care need and patient satisfaction at a hospital emergency department. Health Care Manag Rev. 1992;17:71–5 PMID: 1601663.
8. Watson WT, Marshall ES, Fosbinder D. Elderly patients' perceptions of care in the emergency department. J Emerg Nurs. 1999;25:88–92 PMID: 10097197.
9. Thompson DA, Yarnold PR, Williams DR, Adams SL. Effects of actual waiting time, perceived waiting time, information delivery, and expressive quality on patient satisfaction in the emergency department. Ann Emerg Med. 1996;28:657–65 PMID: 8953956.
10. Antonides G, Verhoef PC, van Aalst M. Consumer perception and evaluation of waiting time: a field experiment. J Consum Psychol. 2002; 12:193–202.
11. Katz KL, Larson BM, Larson RC. Prescription for the waiting in line blues: entertain, enlighten, engage. MIT Sloan Manag Rev. 1991;32:44–53.
12. Pruyn A, Smidts A. Effects of waiting on the satisfaction with the service: beyond objective time measures. Int J Res Mark. 1998;15: 321–34.
13. Davenport PJ, O'Connor SJ, Szychowski JM, Landry AY, Hernandez SR. The relationship between emergency department wait times and inpatient satisfaction. Health Mark Q. Taylor & Francis; 2017;34(2):97–112.

14. Mayer TA, Zimmermann PG. ED customer satisfaction survival skills: one hospital's experience. J Emerg Nurs. 1999;25:187–91.
15. Soremekun OA, Takayesu JK, Bohan SJ. Framework for analyzing wait times and other factors that impact patient satisfaction in the emergency department. J Emerg Med. 2011;41:686–92 PMID: 21440402.
16. Di Somma S, Paladino L, Vaughan L, Lalle I, Magrini L, Magnanti M. Overcrowding in emergency department: an international issue. Intern Emerg Med. 2015;10:171–5 PMID: 25446540.
17. Toloo GS, Aitken P, Crilly J, FitzGerald G. Agreement between triage category and patient's perception of priority in emergency departments. Scand J Trauma Resusc Emerg Med. 2016;24:1–8 PMID: 27756416.
18. Muntlin Å, Gunningberg L, Carlsson M. Patients' perceptions of quality of care at an emergency department and identification of areas for quality improvement. J Clin Nurs. 2006;15:1045–56 PMID: 16879549.
19. Reader TW, Gillespie A, Roberts J. Patient complaints in healthcare systems: a systematic review and coding taxonomy. BMJ Qual Saf. 2014; 23:678–89.
20. Wong LL, Ooi SBS, Goh LG. Patients' complaints in a hospital emergency department in Singapore. Singap Med J. 2007;48:990–5.
21. Sun BC, Adams JG, Burstin HR. Validating a model of patient satisfaction with emergency care. Ann Emerg Med. 2001;38:527–32 PMID: 11679864.
22. Sun BC, Adams J, Orav EJ, Rucker DW, Brennan TA, Burstin HR. Determinants of patient satisfaction and willingness to return with emergency care. Ann Emerg Med. 2000;35:426–34 PMID: 10783404.
23. Trout A, Magnusson AR, Hedges JR. Patient satisfaction investigations and the emergency department: what does the literature say? Acad Emerg Med. 2000;7:695–709 PMID: 10905652.
24. Boudreaux ED, Mandry CV, Wood K. Patient satisfaction data as a quality indicator: a tale of two emergency departments. Acad Emerg Med. 2003;10: 261–8 PMID: 12615592.
25. Hedges JR, Trout A, Magnusson AR. Satisfied patients exiting the emergency department (SPEED) study. Acad Emerg Med. 2002;9:15–21 PMID: 11772664.
26. Krishel S, Baraff LJ. Effect of emergency department information on patient satisfaction. Ann Emerg Med. 1993;22:568–72 PMID: 8442546.
27. Boudreaux ED, O'Hea EL. Patient satisfaction in the emergency department: a review of the literature and implications for practice. J Emerg Med. 2004;26:13–26.
28. Toma G, Triner W, McNutt L-A. Patient satisfaction as a function of emergency department previsit expectations. Ann Emerg Med. 2009;54:360–7 e6. PMID: 19282060.
29. Spahr CD, Flugstad NA, Brousseau DC. The impact of a brief expectation survey on parental satisfaction in the pediatric emergency department. Acad Emerg Med. 2006;13:1280–7 PMID: 17099193.
30. Stata/IC: Stata for mid-sized datasets. StataCorp LLC, College Station, TX. 2019. https://www.stata.com/products/which-stata-is-right-for-me. Accessed 17 June 2019.
31. Kraaijvanger N, Rijpsma D, van Leeuwen H, Edwards M. Self-referrals in the emergency department: reasons why patients attend the emergency department without consulting a general practitioner first—a questionnaire study. Int J Emerg Med. 2015;8:1–6.
32. Toloo S, Fitzgerald G, Aitken P, Ting J, Tippett V, Chu K, et al. Emergency health services: demand and services delivery models. Monograph 1: Literature review and activity trends. Australia: Queensland University of Technology; 2011.
33. Meek R, Phiri W. Australasian triage scale: consumer perspective. Emerg Med Australas. 2005;17:212–7 PMID: 15953221.
34. Gilmartin J, Wright K. Day surgery: patients' felt abandoned during the preoperative wait. J Clin Nurs. 2008;17:2418–25 PMID: 18705721.
35. Shaikh S, Jerrard D, Witting M, Winters M, Brodeur M. How long are patients willing to wait in the emergency separtment before leaving without being seen. West J Emerg Med. 2012;13:463–7 PMID: 23359833.
36. Vieth TL, Rhodes KV. The effect of crowding on access and quality in an academic ED. Am J Emerg Med. 2006;24:787–94 PMID: 17098098.
37. Tran TP, Schutte WP, Muelleman RL, Wadman MC. Provision of clinically based information improves patients' perceived length of stay and satisfaction with EP. Am J Emerg Med. 2002;20(6):506–9.
38. Lin YK, Lin CJ. Factors predicting patients' perception of privacy and satisfaction for emergency care. Emerg Med J. 2011;28:604–8 PMID: 20660898.
39. Bowling A. Mode of questionnaire administration can have serious effects on data quality. J Public Health (Bangkok). 2005;27:281–91 PMID: 15870099.
40. Bujang MA, Sa'at N, Tg Abu Bakar Sidik TMI, Chien Joo L. Sample size guidelines for logistic regression from observational studies with large population: emphasis on the accuracy between statistics and parameters based on real life clinical data. Malaysian J Med Sci. 2018;25:122–30.

Publisher's Note

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

3.2.2 Second article: “Use of a semiautomatic text message system to improve satisfaction with wait time in the adult emergency department: cross-sectional survey study" (2)

We designed this study to measure the patient experience among those who used the semiautomatic message system allowing them to wait outside the ED. A 12-item questionnaire was completed by 100 patients presenting to the HUG outpatient emergency unit, mostly classified as level 3 according to the Swiss Emergency Triage Scale. We also analyzed caregiver acceptance of the system through a questionnaire rating this option with the 21-item Unified Theory of Acceptance and Use of Technology (UTAUT).

Results showed that 97/100 (97%) patients were satisfied with the text message system. Among these, 92/100 (92%) had sufficient time to return to the ED, 95/100 (95%) considered that the message was clear, and 28/100 (28%) described feeling stressed by waiting outside the ED. 75/100 (75%) patients were interested by the system after 30 minutes’ wait time and 87/100 (87%) after 1 hour. The analysis demonstrated that the mean waiting time does not have a statistically significant influence on overall satisfaction.

Based on a 5-point Likert scale ("totally disagree", "partially disagree", "neither agree nor disagree", "partially agree", "totally agree”), we explored if satisfaction with the wait time was influenced by the mean wait time, as described when the patient waits in the waiting room. We performed an analysis of variance (ANOVA) statistical test, which showed no significant differences between wait time means as a function of wait time satisfaction. Conversely, in the analysis of the perception of waiting time when the patient remained in the waiting room, there was a correlation between satisfaction and waiting time (1).

User analysis (nurses) confirmed the ergonomic design of the system. Although most users intended to use it, they did not find it useful. However, since its initial development, we have added six languages and observed that the system is more and more used and concerned approximately 20% of patients consulting the outpatient emergency unit in 2023 (data from HUG Dashboard). In the context of our ED, the system is currently available in eight different languages (Arabic, English, Farsi, French, German, Italian, Spanish and Tigrigna) and is an extension of a similar project developed in the pediatric ED at HUG (11).

From a practical point of view, the nurses decide if the patient can wait outside the ED using the semi-automatic message system and they also decide when the most opportune time is to return. They work with a specific dashboard (Figure 5) permitting to monitor the patients waiting outside. At that time, the patient receives a first message to say that the medical doctor will see him/her in 20 minutes, followed by a reminder message four times every 5 minutes; the final message will inform the patient that he/she will lose the allocated slot in the ED in the case of non-presentation.

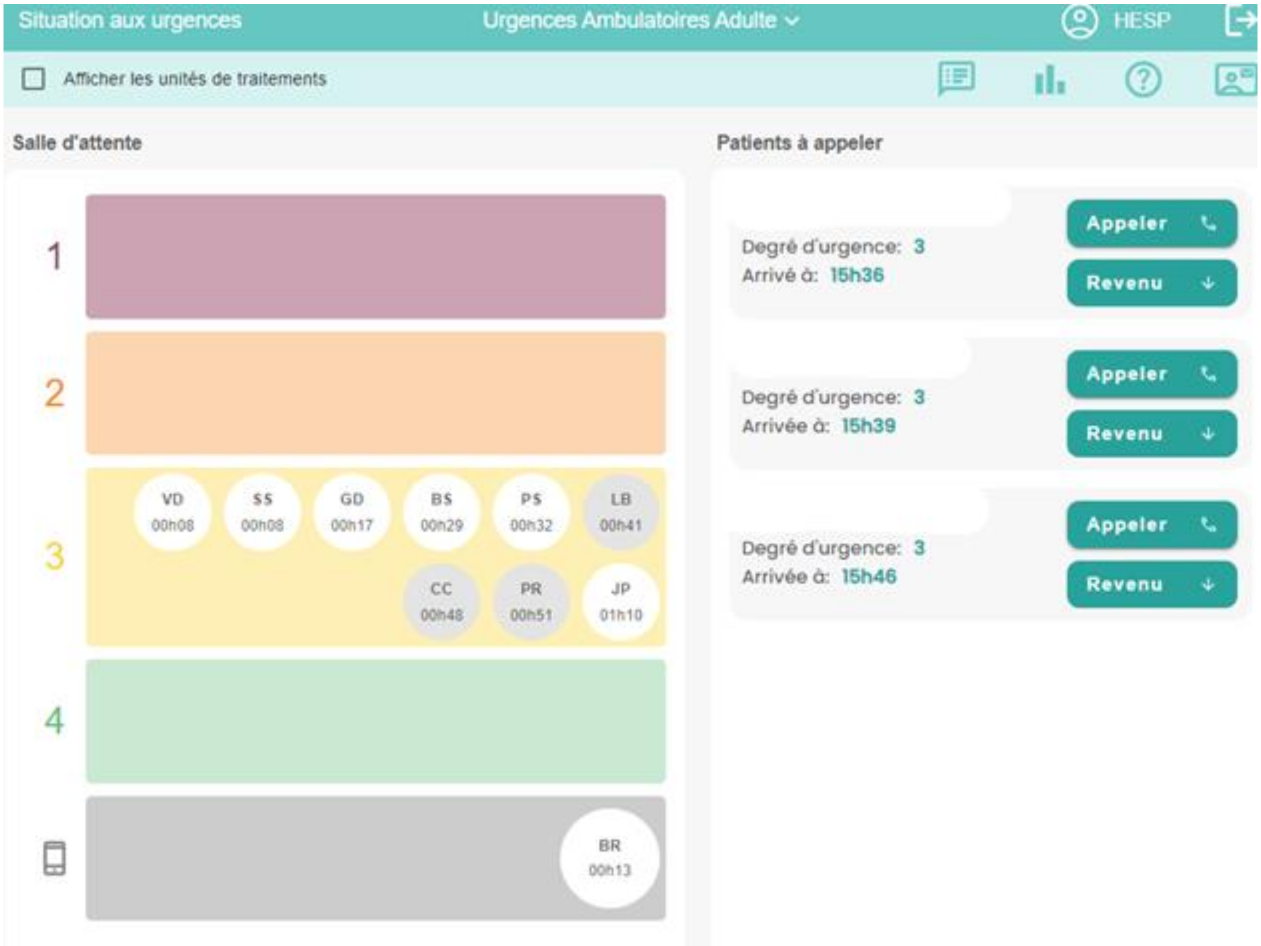


Figure 5: Dashboard for monitoring the patients waiting outside the ED. (Reproduced from the application Infomed in the DPI (Dossier Patient Informatique [electronic patient record]) at HUG.

Original Paper

Use of a Semiautomatic Text Message System to Improve Satisfaction With Wait Time in the Adult Emergency Department: Cross-sectional Survey Study

Frederic Ehrler[1,2], PhD; Jessica Rochat[1,2], MSc; Johan N Siebert[2,3], MD; Idris Guessous[4,5], MD, PhD; Christian Lovis[1,2], MPH, MD; Hervé Spechbach[2,6], MD

[1]Division of Medical Information Sciences, University Hospitals of Geneva, Geneva, Switzerland
[2]Faculty of Medicine, University of Geneva, Geneva, Switzerland
[3]Department of Pediatric Emergency Medicine, Geneva Children's Hospital, Geneva University Hospitals, Geneva, Switzerland
[4]Division of Primary Care, Geneva University Hospitals, Geneva, Switzerland
[5]Department of Health and Community Medicine, Faculty of Medicine, University of Geneva, Geneva, Switzerland
[6]Ambulatory Emergency Care Unit, Department of Primary Care Medicine, Geneva University Hospitals, Geneva, Switzerland

Corresponding Author:
Frederic Ehrler, PhD
Division of Medical Information Sciences
University Hospitals of Geneva
Gabrielle Perret Gentil 4
Geneva, 1205
Switzerland
Phone: 41 79 553 16 03
Email: frederic.ehrler@hcuge.ch

Abstract

Background: Many factors influence patient satisfaction during an emergency department (ED) visit, but the perception of wait time plays a central role. A long wait time in the waiting room increases the risk of hospital-acquired infection, as well as the risk of a patient leaving before being seen by a physician, particularly those with a lower level of urgency who may have to wait for a longer time.

Objective: We aimed to improve the perception of wait time through the implementation of a semiautomatic SMS text message system that allows patients to wait outside the hospital and facilitates the recall of patients closer to the scheduled time of meeting with the physician.

Methods: We performed a cross-sectional survey to evaluate the system using a tailored questionnaire to assess the patient perspective and the Unified Theory of Acceptance and Use of Technology questionnaire for the caregiver perspective. We also monitored the frequency of system use with logs.

Results: A total of 110 usable responses were collected (100 patients and 10 caregivers). Findings revealed that 97 of 100 (97%) patients were satisfied, with most patients waiting outside the ED but inside the hospital. The caregiver evaluation showed that it was very easy to use, but the adoption of the system was more problematic because of the perceived additional workload associated with its use.

Conclusions: Although not suitable for all patients, our system allows those who have a low-severity condition to wait outside the waiting room and to be recalled according to the dedicated time defined in the Swiss Emergency Triage Scale. It not only has the potential to reduce the risk of hospital-acquired infection but also can enhance the patient experience; additionally, it was perceived as a real improvement. Further automation of the system needs to be explored to reduce caregiver workload and increase its use.

(JMIR Med Inform 2022;10(9):e34488) doi: 10.2196/34488

KEYWORDS

emergency; patient satisfaction; service-oriented health care; quality of care; health service; emergency department

Introduction

Background

Patients triaged with low priority in the emergency department (ED) are likely to have a long wait before being seen by a physician, as those with life-threatening and serious conditions are prioritized over patients that are less acute [1]. A side effect of long wait times is the risk that patients leave the ED without being seen by a physician, with this risk increasing significantly after a 1-hour wait [2]. It has also been shown that long wait times can result in staff interruptions by frustrated patients and lead to violent behavior [3,4]. Additionally, it has been reported that a long wait time increases the risk of contracting hospital-acquired infections [5]. As an example, Beggs et al [6] showed that the number of new cases of airborne infections increased substantially with time spent in the waiting room and the number of people waiting. However, reducing overcrowding in the ED waiting room is not a simple task [7]. The space available is often limited and the nature of the ED does not allow for a control on its occupation, which varies significantly over the course of a single day [8,9].

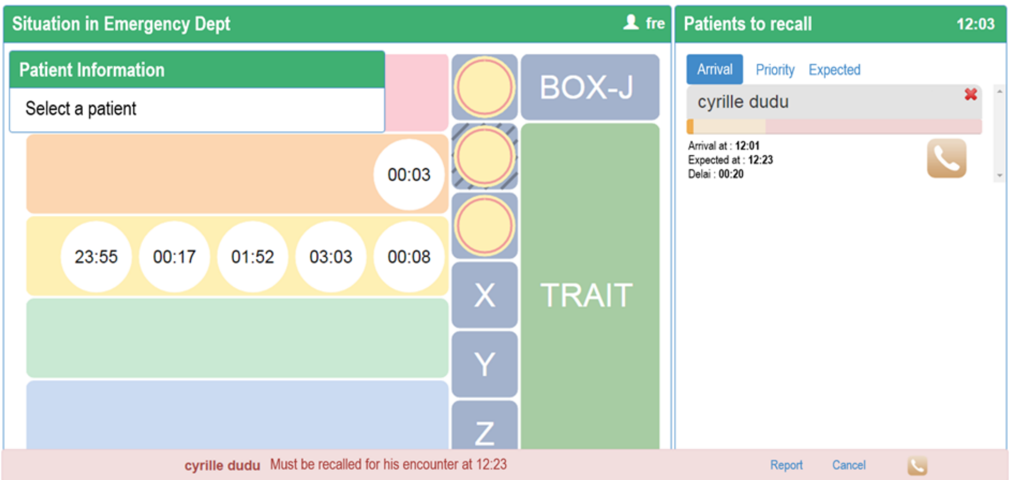
Several attempts have been performed to improve the wait time experience in the ED, either by minimizing the duration between triage and patient care or by acting on the actual perception of wait time [7,10]. Although organizational measures can improve ED efficiency, such as fast track [11], improved triage [12], and better team communication, they will never prevent overcrowding situations, as ED staff cannot be adjusted as quickly as the influx evolves. By contrast, improving the patient experience during the wait has been favored through interventions such as providing information to the patient about the expected wait duration [13], comfort improvement in the waiting room [14], or giving a pager to patients, which allowed them to wait in a place other than the waiting room [15]. These interventions were shown to have a positive impact and are promising strategies to be further explored.

To reduce ED congestion and improve the perception of wait time, we developed a semiautomated message (SMS text message) system that allows patients to wait outside the emergency waiting room and to be recalled closer to the actual time of the medical consultation. In this study, we explore the perceptions of this system by patients and caregivers.

A Semiautomatic SMS Text Message System

The system was initially developed at the pediatric department of Geneva University Hospitals (Geneva, Switzerland) [16], adapted later for the adult setting and deployed in September 2017, and finally introduced in the gynecology and obstetrics setting in 2019. It aims at improving patient flow in EDs by providing caregivers with a system to monitor the flow and ED occupancy. The system allows triaged patients with a low-severity grade to wait outside the ED and to be called back by a recall SMS text message system shortly before they are to be seen by a physician. A screen available to nurses provides real-time occupancy of the emergency rooms and wait times by triage level (Figure 1).

Figure 1. Main screen of the SMS text message recall system. The left-hand side represents the waiting queue in the emergency department waiting room, with each line representing the emergency level and each circle a patient currently waiting. The vertical middle row represents the emergency rooms and their occupancy, with each patient also represented by a circle. The right-hand side is the SMS text message recall system. The patients enrolled are presented with information on their arrival time and expected meeting time with the physician.



Once triaged, each patient can be registered in the SMS text message system by a nurse. A screen displays the patient's key administrative information, allowing the administrative clerk to verify the validity of the patient's telephone number. The nurse estimates the length of the wait and validates the patient's registration in the system. The patient then receives a confirmation message and can leave the ED while remaining virtually in the queue. Whether they are physically present or not, all patients are moved forward normally in the queue and recalled based on their arrival time and emergency level. All registered patients waiting outside the ED are visible on a screen with a time bar individually associated with them and showing the expected time to being seen by a physician. The time bar progressively changes color based on the elapsed time and actions that need to be taken by the caregiver responsible for calling the patient back. A green bar indicates that no recall is needed yet since the meeting with the physician is still distant. The bar turns orange 20 minutes before the patient's scheduled return, suggesting that the triage nurse call the patient back, without being mandatory. If the scheduled return time has passed, the bar turns red.

Dispatch of the first recall SMS text message is left to the discretion of the triage nurse to determine the most opportune time to return for the visit. If the patient does not arrive within 20 minutes of the first SMS text message, the system is automated to send reminder messages every 20 minutes (total of four SMS text messages). At any time, the nurse has the possibility to inform the patient about the evolution of the situation at the ED by sending a predefined message to the patient. Depending on the situation, the message sent will inform the patient that the visit is postponed due to a strong influx of patients or that an early return is possible due to an improved situation. If the patients do not arrive after three reminder SMS text messages, a final fourth SMS text message is sent to inform them that the position in the queue is no longer guaranteed, but the visit is still possible.

Methods

Study Design

This study is a cross-sectional descriptive investigation using a mixed methods methodology, including an assessment of the patients' experience during their wait in the ED through a tailor-made questionnaire, analysis of the system log to understand the use trend, and an assessment of nurses' acceptance of the SMS text message recall system. The survey was conducted between March 13 and April 28, 2017, at the 24-hour ED outpatient unit at Geneva University Hospitals, the largest public hospital in Switzerland with 70,000 patient admissions each year. Utilization logs were collected from October 2017 to August 2019.

ED Setting: Emergency Outpatient Unit

Medical and traumatic pathologies are treated in 12 consultation rooms. Patients wait in a semienclosed waiting room with seating, a television, water, and newspapers. The staff (clinicians and nurses) are the same for the entire unit. The median length of stay is 3.5 hours, with a median waiting time of 1.5 hours.

When patients arrive at the main ED entrance, they are first seen by a triage nurse who decides whether the patient is a candidate for the outpatient unit, based on the Swiss Emergency Triage Scale [17]. Level 1 is a life-/limb-threatening situation where the patient must be seen immediately by a physician. Level 2 must be seen within 20 minutes, level 3 within 120 minutes, and level 4 is considered nonurgent. A total of 80% of patients who come to the ED are classified as level 3 and 10% as level 4. After triage, the patient goes through an administrative registration process and is then directed to one of the subunits by following colored lines on the floor. These lead to a nurse's desk where a nurse escorts the patient to the waiting room. Whenever possible, the nurses inform patients of the estimated waiting time. As soon as a consultation room and physician are available, the patient is taken to the room by the nurse. After the medical visit, the patient can either go home or may have to undergo an additional examination and return to the waiting room. A small percentage of patients (8%) are hospitalized and 5% leave the unit without being seen by a physician [18]. The probability of leaving the ED prematurely is linked to flow concern as demonstrated in previous studies [19].

Study Participants

Patients presenting to the ED outpatient unit with a triage level of 3 and 4 (according to the 4-level Swiss Triage Scale) were invited to participate in the questionnaire part of the study if they were 16 years of age or older and spoke French. We used a convenience sampling method and arbitrarily defined the sample size as 100 participants. Exclusion criteria were patients not capable of discernment (eg, unconscious, intoxicated, extreme trauma, or cognitive impairment), unable to read/understand French, vision problems, severe pain or overly aggressive, and those who had already completed the questionnaire.

Measurement Instruments

Patient Satisfaction Questionnaire

A 12-item questionnaire was designed to assess the patient experience among those who had used the SMS text message system. This questionnaire was of our own design. It contained an item aiming to determine the minimum expected wait time before patients find the system useful. It also explored where the patient waited until being taken care of, whether the advertised waiting time matched the actual waiting time, and whether the content of the SMS text message was clear. Users were asked if they felt stressed during the wait, if they had enough time to come back to the emergency room, and if they were satisfied with the system overall. In addition, the actual wait time for each patient who completed the questionnaire was extracted from the hospital clinical information system.

Caregiver Acceptance Questionnaire

The 21-item Unified Theory of Acceptance and Use of Technology (UTAUT) questionnaire is a unified technology acceptance model formulated by Venkatesh et al [20] as a conceptual framework to understand users' intended use and acceptance of new information technologies, which can be determined by 5 constructs: (1) performance efficiency (4 questions), (2) effort expectancy (4 questions), (3) social influence (4 questions), (4) facilitating conditions (4 questions), and (5) behavioral intention to use the system in the future (3 questions). Each question is scored on a 7-point Likert scale. The questionnaire was distributed anonymously to all nurses working with the system.

System Use Logs

System use was assessed by analysis of the system use logs. A log, including a time stamp, was generated each time a caregiver entered a patient into the SMS text message system, as well as each time a SMS text message was sent.

Procedure and Ethical Considerations

The Geneva Institutional Ethics Committee approved the study protocol (Réq-2016-00555). Patient participation in the study was voluntary, and oral consent was obtained prior to the intervention. After verification of the inclusion criteria, the nurse asked the patients if they agreed to use the SMS text message recall system. Information about the study and confidentiality were given verbally. If accepted, the patients were allowed to wait wherever they wanted (ie, in or outside the ED). We did not verify where the patient waited as it would have been difficult to trace. We arbitrarily decided to set the number of questionnaires to be completed at 100.

Once back in the ED, the patient was immediately brought to a consultation room. The patient was given the study questionnaire by a nurse while waiting for the physician. The nurse remained available for any questions and to assist the patient in completing the questionnaire if necessary. Instructions were given to the medical staff to see the patients immediately after completion of the questionnaire. Once completed, questionnaires were collected by nurses and placed in a dedicated box in a secure room. Questionnaires were collected each morning by a scientific collaborator, and the responses were entered into an Excel (Microsoft Corporation) file. To link the questionnaire data to data extracted from the hospital clinical information system, we used a mapping file linking the questionnaire ID to the patient ID. Once all data were included in the Excel file, only the questionnaire ID was retained to ensure anonymous analysis.

Statistical Analysis

Descriptive statistics were generated to present the demographic and medical characteristics of participants. The difference of the average mean between each level of patient satisfaction was analyzed using an ANOVA performed on SPSS 26 (IBM Corp) software. The caregiver acceptance questionnaire was analyzed by computing the proportion of each response for a given item. UTAUT scores were reported as the average score given to all items of a given dimension for all participants. System logs were analyzed by looking at the number of SMS text messages sent each month during the observation period.

Results

Demographics

Patient questionnaires were distributed between March 13 and April 28, 2017, by a total of 20 nurses during two different shifts (7:30 AM to 4 PM and 3 PM to 11:30 PM). The total number of collected questionnaires was 100. One patient was excluded because he visited the unit twice during the study period. The questionnaire took an average of 10 minutes to complete. Baseline patient demographics and data related to the medical encounter are shown in Table 1. Of the 100 respondents, 87 (87%) were classified with an emergency level of 3, and 12 (12%) were classified in level 4. No patients were classified as levels 1 or 2 as these acuity triage levels require immediate care. The nurses' questionnaire was proposed to all nurses of the unit (n=25) but was only completed by 10 nurses.

Table 1. Demographics of participants and information on their medical encounter.

	Participants (N=100)
Age (years), mean (SD)	38 (14.75)
Sex, n (%)	
Male	60 (60)
Female	40 (40)
Triage level, n (%)	
3	87 (87)
4	12 (12)
Missing	1 (1)
Wait time (hours), n (%)	
<1	32 (32)
1-2	45 (45)
2-3	14 (14)
3-4	8 (8)
>4	1 (1)

Patient Satisfaction Questionnaire

As presented in Table 2, of the total 100 respondents, 97% (n=97) were satisfied with the SMS text message system. Among these, approximately 75% (n=75) were totally satisfied with their waiting time and 56% (n=56) were satisfied. A total of 79 (79%) respondents waited outside of the ED but inside the hospital, as the facility offers the possibility to wait in pleasant places such as the cafeteria, adjacent green spaces, and the meditation room. The fact that patients waited close to the ED was confirmed by the fact that 86% (n=86) of patients returned to the ED on foot. Therefore, 92 (92%) patients had sufficient time to return to the ED once recalled. A total of 95 (95%) patients considered the SMS text message to be clear and 72 (72%) did not feel particularly stressed waiting outside the ED.

Table 2. Questionnaire results.

	Participants (N=100), n (%)
Where did you spend your time while waiting?	
At home	2 (2)
Outside the hospital	13 (13)
Inside the hospital	80 (80)
Other	6 (6)
How do you rate your actual wait time compared to the wait time announced by the nurses?	
Longer	25 (25)
Shorter	49 (49)
Equal	25 (25)
Not informed	1 (1)
The SMS text message content was clearly understandable?	
Totally agree	72 (72)
Partly agree	23 (23)
Neither agree nor disagree	4 (4)
Partly disagree	1 (1)
Totally disagree	0 (0)
Did you experience a feeling of stress linked to your absence from the emergency waiting room?	
Totally agree	8 (8)
Partly agree	10 (10)
Neither agree nor disagree	11 (11)
Partly disagree	22 (22)
Totally disagree	50 (50)
Did you have enough time to return to the emergency room after receiving the recall message?	
Totally agree	59 (59)
Partly agree	33 (33)
Neither agree nor disagree	4 (4)
Partly disagree	2 (2)
Totally disagree	0 (0)
How did you return to the emergency room after receiving the recall message?	
On foot	86 (86)
Public transport	8 (8)
Private transport	2 (2)
Other	4 (4)
Are you satisfied with the SMS text message recall service?	
Totally agree	75 (75)
Partly agree	22 (22)
Neither agree nor disagree	3 (3)
Partly disagree	0 (0)
Totally disagree	0 (0)
Were you satisfied with your waiting time?	
Totally agree	28 (28)
Partly agree	28 (28)

	Participants (N=100), n (%)
Neither agree nor disagree	20 (20)
Partly disagree	12 (12)
Totally disagree	11 (11)

By asking patients what would be the minimum duration of expected wait that would trigger an interest to be enrolled in the system in a future encounter (Table 3), we found that 45 of the 100 (45%) patients were interested in the system regardless of the waiting time. After 30 minutes of expected waiting time, 75 (75%) patients were interested in the system, and 87 (87%) patients were interested after 1 hour.

Table 3. Patients interested in using the SMS text message system after n minutes.

Number of minutes	Patients, n (%)
0	45 (45)
10	51 (51)
20	63 (63)
30	75 (75)
40	75 (75)
50	75 (75)
60	87 (87)
70	87 (87)
80	88 (88)
90	93 (93)
100	93 (93)
110	93 (93)
120	100 (100)

Satisfaction and Waiting Time

To determine whether wait time duration influenced the level of patient satisfaction with wait time, we assessed if the differences in mean wait time across the five wait time satisfaction modalities (ie, totally disagree, partly disagree, neither agree nor disagree, partly agree, and totally agree) were statistically significant (descriptive statistics are presented in Table 4). As the homogeneity of variance using Levene was not statistically significant (P=.42), meaning that the variances were equal across groups, an ANOVA was performed. We found no significant differences between wait time means as a function of wait time satisfaction (P=.32; F_4=1.193).

Table 4. Average wait duration according to user satisfaction with wait time.

Satisfaction with wait time	Wait time (min), mean (SD)	Participants (N=100), n (%)
Totally disagree	86.9091 (64.28)	11 (11)
Partly disagree	105.0833 (58.78)	12 (12)
Neither agree nor disagree	101.5000 (58.06)	20 (20)
Partly agree	98.3929 (47.88)	28 (28)
Totally agree	75.0357 (46.22)	28 (28)
Total	91.9495 (53.10)	99 (99)

Caregiver Acceptance Questionnaire

The UTAUT questionnaire distributed to all nurses using the system was completed by 10 nurses (20% participation rate; Table 5). Nurses emphasized the good ergonomics of the system as they rated effort expectancy with an average score of 6.0. This was also confirmed by the facilitating condition dimension, including the resources and knowledge necessary to use the system, which were ranked above 5. Behavioral intention was high as most users intended to use the system frequently in the future on a daily basis. The expected gain on performance was less obvious for respondents. Although most users found the system useful (mean 4.5, SD 1.9), they did not find that the system increased their productivity (mean 3.2, SD 1.6) or speed at work (mean 3.0, SD 1.4). Hedonic motivation ranked below 4 as users did not find the system enjoyable or fun to use. Finally, social influence scored the lowest (mean 2.3, SD 1.9) as all users did not observe a positive influence on their peers or hierarchy toward the use of the system.

Table 5. Score distribution for each UTAUT dimension.

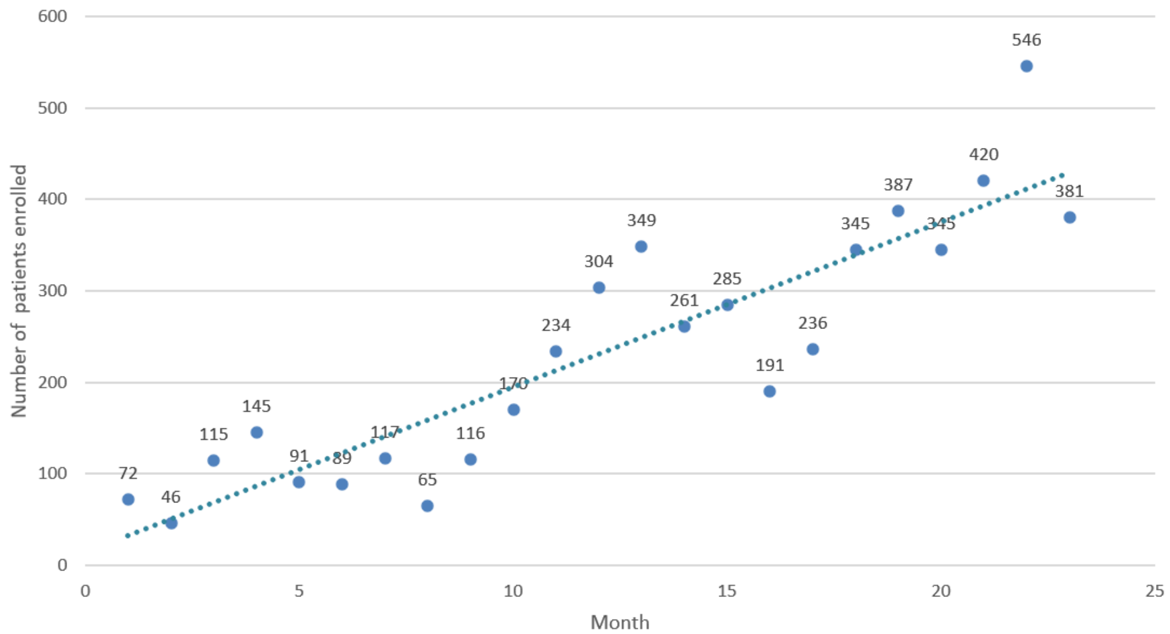
UTAUT[a] dimension	Nurses' scores, n (%)							Score, mean (SD)
	1	2	3	4	5	6	7	
Performance expectancy (n=33)	1 (3)	13 (39)	4 (12)	4 (12)	5 (15)	4 (12)	2 (6)	3.6 (1.8)
Effort expectancy (n=44)	0 (0)	3 (7)	0 (0)	3 (7)	6 (14)	9 (20)	23 (52)	6.0 (1.4)
Social influence (n=19)	12 (63)	0 (0)	0 (0)	4 (21)	2 (11)	0 (0)	1 (5)	2.3 (1.9)
Facilitating condition (n=41)	1 (2)	3 (7)	1 (2)	6 (15)	11 (27)	6 (15)	13 (32)	5.3 (1.6)
Hedonic motivation (n=24)	5 (21)	5 (21)	3 (13)	4 (17)	2 (8)	2 (8)	3 (13)	3.5 (2.1)
Behavioral intention (n=30)	0 (0)	2 (7)	3 (10)	4 (13)	6 (20)	6 (20)	9 (30)	5.3 (1.6)

[a]UTAUT: Unified Theory of Acceptance and Use of Technology.

Log Analysis

Figure 2 shows the number of unique patients entered into the SMS text message system from its introduction on October 1, 2017, to August 31, 2019. Although not always continuous, there was a clear trend of an increase in system use over time, ranging from 46 patients enrolled in November 2017 to 546 in July 2019. This corresponds to the trend of a linear function (18*x + 14) meaning that each month 18 additional patients are included in the system.

Figure 2. Number of unique patients enrolled in the SMS text message system each month (October 1, 2017, to August 31, 2019) and linear trend.



Discussion

Principal Findings

In this study, we found that 87 of the 100 (87%) patients with low-to-moderate urgency were interested in waiting outside the ED waiting room when the expected wait time was 1 hour or more. In a previous study, we observed that patients perceived the wait to be acceptable if it did not exceed 1 hour [21]. After 2 hours, they preferred to leave the ED before seeing a physician [22]. We observed that waiting outside the emergency room was perceived as a source of stress for <20% of participants, possibly related to the perceived reduced control over the situation when outside the room. Indeed, patients waiting outside the waiting room have no view on the current situation and can easily imagine being forgotten by ED staff [23]. Patients may also be concerned that their condition may worsen [24]. Thus, it may be worthwhile to send a recall SMS text message at regular intervals to indicate the patient's current place in the waiting queue to provide reassurance about their position and the progression of the ED process [25]. The messages could also inform the patient about the proper actions in case their condition worsens, such as approaching a specific person or to go to a specific desk to advise staff. This type of concern has already been highlighted in another report showing that some patients want to remain visible to the caregiver to avoid being forgotten [26].

By comparing the relationship between wait time and satisfaction with our previous study [21] performed in a similar setting, we observed a reduced negative influence of the average wait time on patient satisfaction. Whereas in previous studies longer wait times led to significantly less satisfied patients, this relationship was no longer observed in our study [27]. This may indicate that patients are less concerned about the length of wait if they can wait in another location than in a waiting room where they have little to do but remain seated until they are taken care of. This correlates well with our results indicating that most patients were willing to use the system if the wait was longer than 1 hour.

Use of the system by the nursing staff began at a low frequency but increased steadily over time. Nurses' initial reaction to the system was negative or neutral, and they initially perceived the tool as an additional burden to their workload. This phenomenon has already been observed in other studies [28,29]. The use of the system by many patients allowed it to predict potential benefits of the tool, such as reduced interruptions due to inpatient patients and reduced aggressive behavior in the waiting room due to long wait times [30]. However, informal feedback from nurses using the system highlighted the difficulty of using it when the ED was crowded. This is probably because busy nurses have less time to use the system in addition to regular duties that results in a contradictory effect that prevents the system from being used when it would be most useful. There are two solutions that can be considered to deal with this problem. Either the system can be used by administrative staff or the system can be automated. At our institution, the drive to develop this system has been a top-down process, and we plan to employ administrative workers to offload these tasks from nursing staff.

Limitations

A strong limitation of the paper is the absence of a strict control group. To compare the effect of our intervention on the relationship between wait time and patient satisfaction, we used the results of a previous study [21] conducted in the same setting where we explored the factors associated with wait perception as a preintervention finding. However, since the questionnaire was not the same, the comparison is limited. The selection of patients based on their interest in using the SMS text message system must be taken into account as it certainly has an impact on the high satisfaction rate, as well as on the low-stress rate related to a wait outside the ED. Indeed, a patient with a high-stress level could refuse to use the system.

Another limitation is the use of a questionnaire of our own design. Since the questionnaire has not been scientifically validated, we cannot guarantee that it measured accurately the investigated constructs. Unfortunately, we did not record the acceptance rate of the system, and it would have been interesting to see how many patients refused the system and preferred to stay in the waiting room. The low participation rate of nurses is also a limitation, and it will be useful to conduct a further survey following the training of administrative staff to take over tasks.

Conclusions

Waiting in the emergency waiting room is a source of frustration for the patient. In addition to the increase in aggressive attitudes in some patients when the ED waiting room is crowded, it also puts patients at risk of hospital-acquired infections. We observed a high level of satisfaction with our SMS text message recall system, allowing a wait outside the ED, but the adoption was more difficult among nurses. Relying on further automation of the system may be an interesting solution to reduce caregiver workload, but this must be done with caution given the high unpredictability of the ED waiting process.

Acknowledgments

We are grateful to Rosemary Sudan (freelance technical editor) for providing editorial assistance.

Authors' Contributions

All authors contributed to the conception, design, analysis, and interpretation of data for this work. All authors contributed to drafting, revising, and approving the final version of the manuscript.

Conflicts of Interest

JNS, FE, and CL have individual intellectual property rights on the SMS text message recall system and, as employees of Geneva University Hospitals, might receive indirect institutional reward in case of commercialization. The other authors have no conflicts of interest to declare.

References

1. Farrohknia N, Castrén M, Ehrenberg A, Lind L, Oredsson S, Jonsson H, et al. Emergency department triage scales and their components: a systematic review of the scientific evidence. Scand J Trauma Resusc Emerg Med 2011 Jun 30;19:42 [FREE Full text] [doi: 10.1186/1757-7241-19-42] [Medline: 21718476]
2. Li D, Brennan J, Kreshak A, Castillo E, Vilke G. Patients who leave the emergency department without being seen and their follow-up behavior: a retrospective descriptive analysis. J Emerg Med 2019 Jul;57(1):106-113. [doi: 10.1016/j.jemermed.2019.03.051] [Medline: 31078346]

3. Cohen EL, Wilkin HA, Tannebaum M, Plew MS, Haley LL. When patients are impatient: the communication strategies utilized by emergency department employees to manage patients frustrated by wait times. Health Commun 2013;28(3):275-285. [doi: 10.1080/10410236.2012.680948] [Medline: 22716025]
4. Morphet J, Griffiths D, Plummer V, Innes K, Fairhall R, Beattie J. At the crossroads of violence and aggression in the emergency department: perspectives of Australian emergency nurses. Aust Health Rev 2014 May;38(2):194-201. [doi: 10.1071/AH13189] [Medline: 24670224]
5. Quach C, McArthur M, McGeer A, Li L, Simor A, Dionne M, et al. Risk of infection following a visit to the emergency department: a cohort study. CMAJ 2012 Mar 06;184(4):E232-E239 [FREE Full text] [doi: 10.1503/cmaj.110372] [Medline: 22271915]
6. Beggs CB, Shepherd SJ, Kerr KG. Potential for airborne transmission of infection in the waiting areas of healthcare premises: stochastic analysis using a Monte Carlo model. BMC Infect Dis 2010 Aug 20;10:247 [FREE Full text] [doi: 10.1186/1471-2334-10-247] [Medline: 20727178]
7. Yen K, Gorelick MH. Strategies to improve flow in the pediatric emergency department. Pediatr Emerg Care 2007 Oct;23(10):745-9; quiz 750. [doi: 10.1097/PEC.0b013e3181568efe] [Medline: 18090112]
8. Otsuki H, Murakami Y, Fujino K, Matsumura K, Eguchi Y. Analysis of seasonal differences in emergency department attendance in Shiga Prefecture, Japan between 2007 and 2010. Acute Med Surg 2016 Apr;3(2):74-80. [doi: 10.1002/ams2.140] [Medline: 29123756]
9. Downing A, Wilson R. Temporal and demographic variations in attendance at accident and emergency departments. Emerg Med J 2002 Nov;19(6):531-535 [FREE Full text] [doi: 10.1136/emj.19.6.531] [Medline: 12421778]
10. Yuzeng S, Hui LL. Improving the wait time to triage at the emergency department. BMJ Open Qual 2020 Feb;9(1):e000708 [FREE Full text] [doi: 10.1136/bmjoq-2019-000708] [Medline: 32019749]
11. Hampers LC, Cha S, Gutglass DJ, Binns HJ, Krug SE. Fast track and the pediatric emergency department: resource utilization and patients outcomes. Acad Emerg Med 1999 Nov;6(11):1153-1159 [FREE Full text] [doi: 10.1111/j.1553-2712.1999.tb00119.x] [Medline: 10569389]
12. Sun BC, Adams J, Orav EJ, Rucker DW, Brennan TA, Burstin HR. Determinants of patient satisfaction and willingness to return with emergency care. Ann Emerg Med 2000 May;35(5):426-434. [Medline: 10783404]
13. Xie B, Youash S. The effects of publishing emergency department wait time on patient utilization patterns in a community with two emergency department sites: a retrospective, quasi-experiment design. Int J Emerg Med 2011 Jun 14;4(1):29. [doi: 10.1186/1865-1380-4-29] [Medline: 21672236]
14. Goodarzi H, Javadzadeh H, Hassanpour K. Assessing the physical environment of emergency departments. Trauma Mon 2015 Nov;20(4):e23734 [FREE Full text] [doi: 10.5812/traumamon.23734] [Medline: 26839860]
15. Scolnik D, Matthews P, Caulfeild J, Williams C, Feldman B. Pagers in a busy paediatric emergency waiting room: a randomized controlled trial. Paediatr Child Health 2003 Sep;8(7):422-426 [FREE Full text] [doi: 10.1093/pch/8.7.422] [Medline: 20019948]
16. Ehrler F, Lovis C, Rochat J, Schneider F, Gervaix A, Galetto-Lacour A, et al. [InfoKids: changing the patients' journey paradigm in an Emergency Department]. Rev Med Suisse 2018 Sep 05;14(617):1538-1542. [Medline: 30226668]
17. Rutschmann O, Hugli O, Marti C, Grosgurin O, Geissbuhler A, Kossovsky M, et al. Reliability of the revised Swiss Emergency Triage Scale: a computer simulation study. Eur J Emerg Med 2018 Aug;25(4):264-269 [FREE Full text] [doi: 10.1097/MEJ.0000000000000449] [Medline: 28099182]
18. Grosgurin O, Cramer B, Schaller M, Sarasin F, Rutschmann O. Patients leaving the emergency department without being seen by a physician: a retrospective database analysis. Swiss Med Wkly 2013;143:w13889. [doi: 10.4414/smw.2013.13889] [Medline: 24317804]
19. Weiss SJ, Ernst AA, Derlet R, King R, Bair A, Nick TG. Relationship between the National ED Overcrowding Scale and the number of patients who leave without being seen in an academic ED. Am J Emerg Med 2005 May;23(3):288-294. [doi: 10.1016/j.ajem.2005.02.034] [Medline: 15915399]
20. Venkatesh V, Morris M, Davis G, Davis F. User Acceptance of Information Technology: toward a unified view. MIS Q 2003;27(3):425 [FREE Full text] [doi: 10.2307/30036540]
21. Spechbach H, Rochat J, Gaspoz J, Lovis C, Ehrler F. Patients' time perception in the waiting room of an ambulatory emergency unit: a cross-sectional study. BMC Emerg Med 2019 Aug 01;19(1):41 [FREE Full text] [doi: 10.1186/s12873-019-0254-1] [Medline: 31370794]
22. Shaikh S, Jerrard D, Witting M, Winters M, Brodeur M. How long are patients willing to wait in the emergency department before leaving without being seen? West J Emerg Med 2012 Dec;13(6):463-467 [FREE Full text] [doi: 10.5811/westjem.2012.3.6895] [Medline: 23359833]
23. Rochat CP, Gaucher N, Bailey B. Measuring anxiety in the pediatric emergency department. Pediatr Emerg Care 2018 Aug;34(8):558-563. [doi: 10.1097/PEC.0000000000001568] [Medline: 30020249]
24. Dubé L, Teng L, Hawkins J, Kaplow M. Emotions, the neglected side of patient-centered health care management: the case of emergency department patients waiting to see a physician. In: Advances in Health Care Management: Volume 3. Bingley, West Yorkshire: Emerald Publishing Limited; 2002:161-193.

25. Andrade CC, Devlin AS. Stress reduction in the hospital room: applying Ulrich's theory of supportive design. J Environ Psychol 2015 Mar;41:125-134. [doi: 10.1016/j.jenvp.2014.12.001]
26. Yoon J, Sonneveld M. Anxiety of patients in the waiting room of the emergency department. In: Proceedings of the fourth international conference on Tangible, embedded, and embodied interaction. 2010 Jan Presented at: TEI '10; January 24-27, 2010; Cambridge, MA p. 279-286. [doi: 10.1145/1709886.1709946]
27. Davenport P, O'Connor SJ, Szychowski J, Landry A, Hernandez S. The relationship between emergency department wait times and inpatient satisfaction. Health Mark Q 2017;34(2):97-112. [doi: 10.1080/07359683.2017.1307066] [Medline: 28467280]
28. Mareš J. Resistance of health personnel to changes in healthcare. Kontakt 2018 Oct 12;20(3):e262-e272. [doi: 10.1016/j.kontakt.2018.04.002]
29. Safi S, Thiessen T, Schmailzl KJ. Acceptance and resistance of new digital technologies in medicine: qualitative study. JMIR Res Protoc 2018 Dec 04;7(12):e11072 [FREE Full text] [doi: 10.2196/11072] [Medline: 30514693]
30. Efrat-Treister D, Moriah H, Rafaeli A. The effect of waiting on aggressive tendencies toward emergency department staff: Providing information can help but may also backfire. PLoS One 2020;15(1):e0227729 [FREE Full text] [doi: 10.1371/journal.pone.0227729] [Medline: 31995583]

Abbreviations

ED: emergency department
UTAUT: Unified Theory of Acceptance and Use of Technology

Edited by G Eysenbach; submitted 26.10.21; peer-reviewed by A Benis, R Rivera; comments to author 24.11.21; revised version received 14.01.22; accepted 29.04.22; published 06.09.22

Please cite as:
Ehrler F, Rochat J, Siebert JN, Guessous I, Lovis C, Spechbach H
Use of a Semiautomatic Text Message System to Improve Satisfaction With Wait Time in the Adult Emergency Department: Cross-sectional Survey Study
JMIR Med Inform 2022;10(9):e34488
URL: https://medinform.jmir.org/2022/9/e34488
doi: 10.2196/34488
PMID:

3.3.1 Third article: "A speech-enabled fixed-phrase translator for emergency settings: crossover study" (3)

We conducted a cross-over study in the research laboratory of HUG with 12 French-speaking doctors, including 6 from the outpatient emergency unit and 6 general practitioners in the city of Geneva. Apart from one, all were naïve of the tool BabelDr. The aim was to test the capacity of doctors to reach the correct diagnosis and to assess if the voice recognition component in BabelDr was of added value. They used BabelDr to make a diagnosis for two standardized Arabic-speaking patients. The male patient had a nephritic colic and the female had a cystitis. Following a short introduction to the BabelDr system, each doctor was asked to question both patients and to reach a diagnosis. All consultations were videorecorded in the laboratory. We measured eye movement, use of the computer mouse, questions asked and questions translated, the use of the text for asking questions and the duration of interaction. We also measured user satisfaction on a 5-point Likert scale with 23 questions regarding usability, learnability, appropriateness, the speech component, and the overall usefulness of BabelDr. All medical doctors found the correct diagnosis.

Doctors for both scenarios used more speech than text to interact with patients: 28.5 speech interactions for the cystitis case and 4.5 text interaction. For the renal colic case, this was 36 versus 10, respectively. In a regression model with mixed effect using multivariate analysis to adjust for the session and the scenario, we showed that the percentage of translated speech was negatively associated with the translated text ($p=0.02$). However, if the translated speech increased by 10%, the translated text decreased by 2.4 (95% CI, 0.7-4.2), and was improved for the second session, thus implying that the first session acted as training to be more familiar with the BabelDr tool.

Participants were mostly satisfied, but one-half estimated to be too restricted by the rigidity of the tool and there was a negative association between the failure of vocal recognition and satisfaction. All doctors said that they could conclude with the patients and could imagine using BabelDr on a daily basis in similar conditions with their patients. Most considered that the system easily recognized their voice. At the end of the study, we hypothesized that the patient would prefer to see a doctor speaking, even if they do not understand the meaning, than a doctor writing some text, which is then translated orally. For this reason, we conducted a real-life test in the HUG ED of oral speech between doctors and patients, with a focus on satisfaction on both sides (4).

Original Paper

A Speech-Enabled Fixed-Phrase Translator for Emergency Settings: Crossover Study

Hervé Spechbach[1], MD; Johanna Gerlach[2], PhD; Sanae Mazouri Karker[1], MD; Nikos Tsourakis[2], PhD; Christophe Combescure[1], PhD; Pierrette Bouillon[2], Prof Dr

[1]Geneva University Hospitals, Genève, Switzerland

[2]Faculté de traduction et d'interprétation, Université de Genève, Genève, Switzerland

Corresponding Author:
Pierrette Bouillon, Prof Dr
Faculté de traduction et d'interprétation
Université de Genève
40, bd du Pont d'Arve
Genève, 1211
Switzerland
Phone: 41 223798701
Email: Pierrette.bouillon@unige.ch

Abstract

Background: In the context of the current refugee crisis, emergency services often have to deal with patients who have no language in common with the staff. As interpreters are not always available, especially in emergency settings, medical personnel rely on alternative solutions such as machine translation, which raises reliability and data confidentiality issues, or medical fixed-phrase translators, which sometimes lack usability. A collaboration between Geneva University Hospitals and Geneva University led to the development of BabelDr, a new type of speech-enabled fixed-phrase translator. Similar to other fixed-phrase translators (such as Medibabble or UniversalDoctor), it relies on a predefined list of pretranslated sentences, but instead of searching for sentences in this list, doctors can freely ask questions.

Objective: This study aimed to assess if a translation tool, such as BabelDr, can be used by doctors to perform diagnostic interviews under emergency conditions and to reach a correct diagnosis. In addition, we aimed to observe how doctors interact with the system using text and speech and to investigate if speech is a useful modality in this context.

Methods: We conducted a crossover study in December 2017 at Geneva University Hospitals with 12 French-speaking doctors (6 doctors working at the outpatient emergency service and 6 general practitioners who also regularly work in this service). They were asked to use the BabelDr tool to diagnose two standardized Arabic-speaking patients (one male and one female). The patients received a priori list of symptoms for the condition they presented with and were instructed to provide a negative or noncommittal answer for all other symptoms during the diagnostic interview. The male patient was standardized for nephritic colic and the female, for cystitis. Doctors used BabelDr as the only means of communication with the patient and were asked to make their diagnosis at the end of the dialogue. The doctors also completed a satisfaction questionnaire.

Results: All doctors were able to reach the correct diagnosis based on the information collected using BabelDr. They all agreed that the system helped them reach a conclusion, even if one-half felt constrained by the tool and some considered that they could not ask enough questions to reach a diagnosis. Overall, participants used more speech than text, thus confirming that speech is an important functionality in this type of tool. There was a negative association (P=.02) between the percentage of successful speech interactions (spoken sentences sent for translation) and the number of translated text items, showing that the doctors used more text when they had no success with speech.

Conclusions: In emergency settings, when no interpreter is available, speech-enabled fixed-phrase translators can be a good alternative to reliably collect information from the patient.

(JMIR Med Inform 2019;7(2):e13167) doi: 10.2196/13167

KEYWORDS

anamnesis; emergencies; tools for translation and interpreting; fixed-phrase translator; speech modality

Introduction

Background

In the context of the current refugee crisis, emergency services are increasingly confronted with patients who have no language in common with staff and may not share the same culture. For example, at Geneva University Hospitals (HUG), 52% of patients are foreigners and 10% speak no French at all. In 2017, the 10 languages for which interpretation services were the most solicited were Tigrinya, Tamil, Albanian, Farsi, Spanish, Somalian, Syrian, Dari, Portuguese, and Arabic (North Africa). Taken together, these languages represent 75% of the interpreting hours at HUG (Geneva University Hospitals, personal communication, 2017).

This language barrier situation is known to pose many safety and ethical problems: It is responsible for increased risks for patients [1] and is very expensive. For example, as reported by Rechel et al in 2003 [2], the United States Institute for Healthcare Advancement estimated that US $73 billion was wasted annually in the United States as a result of communication problems in health care, many of which originate from language differences. Both ethically and legally, hospitals have a duty to offer all patients the same quality of care, including the right to have a dialogue with health professionals.

Different solutions are available for use in emergency settings to address these language barriers, but they all have their drawbacks. Phone-based interpreter services, which are the most common solution, are generally considered adequate, but they are expensive (3 Swiss francs/minute with AOZ Medios, a national interpreting service mandated by the Swiss Federal Office of Public Health), not always available for some languages, and less satisfactory than face-to-face interaction with a physically present interpreter [3]. Asking patients' relatives to translate speech is known to create substantial risks [1]. Machine translation, such as Google Translate, another low-cost solution more commonly used in emergency contexts, is also extremely problematic, as this type of tool has not been developed for medical use. Some recent studies have estimated that nearly 40% of sentences of medical speech translated by Google Translate are mistranslated [4,5]. However, such systems also pose ethical problems and are not currently compatible with the Swiss Data Protection Law. A plethora of specialized systems have also been developed for medical communication, both in the academic and industry settings (including fixed-phrase translation or machine translation systems [6]), but it is not always clear how they were built or evaluated and if they are extensible. As emphasized in the recent review by Dew et al [6], there is a lack of criteria for the development and evaluation of these systems, which impedes the adoption of these systems in emergency settings.

For these reasons, we have developed a new type of speech-enabled fixed-phrase translation tool for medical dialogue (BabelDr [7]), based on our previous experience in the field [8] in a collaborative venture between HUG and the University of Geneva Faculty of Translation and Interpreting. This tool is a compromise between speech-to-speech machine translation and fixed-phrase translation systems and directly addresses specific needs in emergency settings (ie, high accuracy, extensibility, portability to low-resource languages and domains, and data security). It was also designed as a way to collect doctor-patient dialogues and thereby improve our understanding of the criteria for the development of this type of system.

This study is the first step in this direction. It aims to determine whether this type of restricted translation tool can be used by doctors to perform a diagnostic interview and reach a correct diagnosis and to quantify if speech adds value to fixed-phrase translators. Although different evaluations of medical devices have been conducted [6], to the best of our knowledge, this is the first study that attempts to show the impact of "phraselators" on the diagnosis and to define a methodology to achieve this.

The BabelDr App

The BabelDr app can be characterized as a "phraselator" [9,10]. Similar to well-known medical fixed-phrase translation apps such as Medibabble [11] or UniversalDoctor [12], the system relies on a set of predefined sentences (mostly yes/no medical questions or instructions) translated by human translators to ensure translation reliability. However, in contrast to traditional fixed-phrase translators, the doctor can also freely ask his/her question and the system will match the recognition result to the closest predefined sentence in the list. The app was designed from the beginning to meet the hospital's needs. In particular, it is easy to extend it to new target languages and situations in order to follow demographic changes and allow its integration in different services. The content is described efficiently with rules (synchronized grammar [13]) that map multiple synonymous patterns ("variations") to a sentence expressing the core meaning ("core sentence"). For example, "Do you have a fever?" "Is your temperature high?" and "Have you observed a high temperature?" will all be mapped to the core sentence "Do you have fever?" In addition, patterns with variables (eg, "Is it a QUALITATIVE pain?" "Do you have a QUALITATIVE pain?" etc, where "QUALITATIVE" is a variable that can take multiple values such as "severe" and "dull") allow the description of content in a productive way. The system currently contains around 2500 patterns and 600 variables, linking more than one billion variations to approximately 25,000 core sentences. Translation follows the usual standards and is performed online with translation memory in two steps—translation of patterns followed by revision of complete sentences [14]. Target languages focus on the languages important for HUG (Spanish, Arabic, Swiss French sign language, Tigrinya, Farsi, Dari, and Albanian). To ensure data confidentiality, both speech recognition and translation are carried out on secure local servers and all interactions are saved locally.

For speech recognition and matching, the system combines rule-based and robust methods, derived from the rules. When the doctor speaks, the system first recognizes what is said using both a grammar-based version of "Nuance" and a specialized statistical version (Nuance Communications Inc, Burlington, MA). It then maps the recognition results to the closest core sentence using both rules and robust matching techniques borrowed from information retrieval, described in detail elsewhere [15]. This closest core sentence is then translated orally for the patient who will answer nonverbally. As it is not an exact translation of the doctor's question, but a translation of the corresponding matched sentence, the core sentence is always echoed back to the doctor, so that he can verify what the system understood. The translation is thus only produced for the patient if the core sentence is approved by the doctor. Therefore, core sentences play a crucial role in the process by not only providing feedback to doctors concerning recognition accuracy, but also making the meaning of the sentence explicit for both translators and patients [16,17]. These core sentences were designed very carefully with doctors and translators, so that they are as accessible and explicit as possible in order to avoid communication problems. In addition to using core sentences for the verification of translations, users can also access them directly by browsing and searching via keywords. The associated translations can then be submitted without the need for further checking, similar to other phraselators [16].

Figure 1 illustrates the BabelDr interface and how an interaction is carried out. The doctor first selects the diagnostic domain based on the main patient complaint (headache, abdominal pain, dermatological problem, etc) and the language and gender of the patient (male or female). He/she can then speak a sentence ("speech interaction"). If the echoed core sentence corresponds to what the doctor wants to ask, he/she can click on it to produce the translation for the patient. In addition to speech input, doctors can search the list of core sentences using keywords (only with exact matching, as in traditional phraselators) and click on sentences to translate them for the patient ("text translated").

After translating a sentence to the patient (Figure 2), the translation is produced both in text and spoken form. The coverage list is automatically scrolled to the latest core sentence translated, giving quick access to related questions. The translated sentence is also added to a history list that can be downloaded as a PDF at the end of the dialogue.

Figure 1. Screenshot of the BabelDr app.

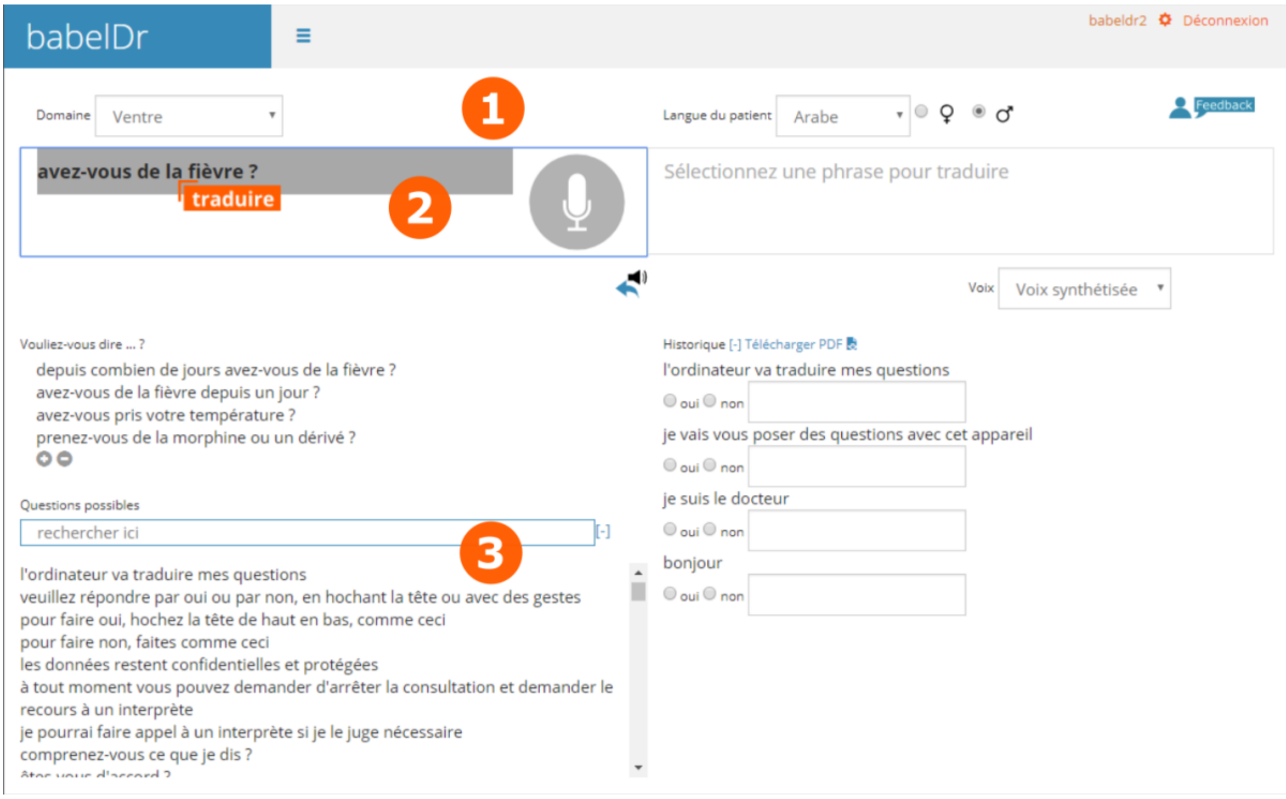
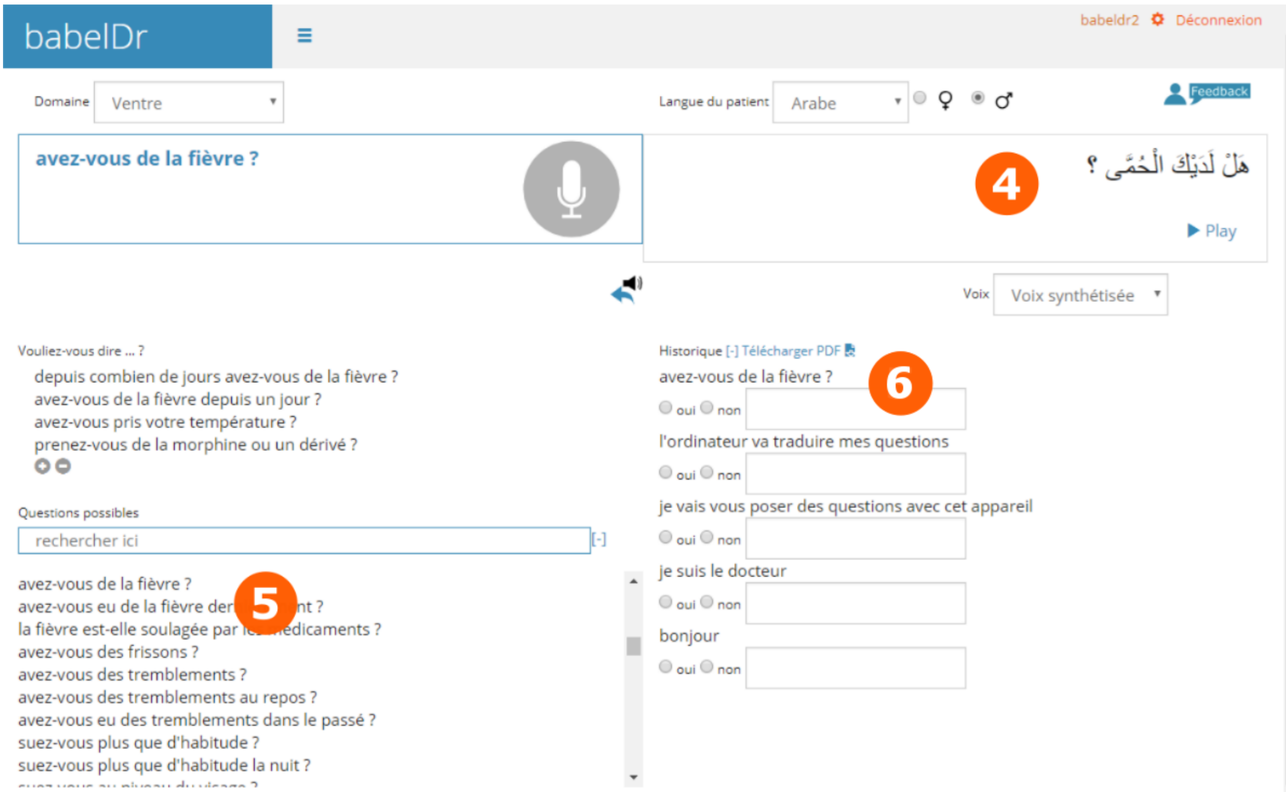


Figure 2. Example of the interface after the doctor translated “Avez-vous de la fièvre?” (Do you have fever?).



Methods

Identifying the Research Questions

This study aims (1) to determine whether a restricted translation tool like BabelDr can be used by doctors to perform a diagnostic interview and reach a correct diagnosis and (2) to quantify how doctors use text versus speech interactions in order to investigate if speech adds value to fixed-phrase translators. Our hypotheses were that this type of tool would demonstrate good functional suitability (doctors can collect all the information necessary to reach a diagnosis in an efficient way) and usability (doctors will use more speech to interact than text, as speaking should allow them to communicate more naturally, like when working with interpreters).

Design

The study was conducted at the HUG research laboratory in December 2017. In this crossover trial, 12 French-speaking doctors were asked to use BabelDr to diagnose two standardized Arabic-speaking patients (one male and one female) whose main complaint was lower back pain. The male patient was standardized for nephritic colic and the female patient, for cystitis. These two diagnoses are among the 10 most frequent at HUG (Geneva University Hospitals, personal communication, 2018). Each of the 12 doctors carried out a diagnostic interview with both patients, where half of the doctors began with the male patient and the other half began with the female patient.

Before the diagnostic interviews, doctors were informed about the main patient complaint (pain in the lower back). At the beginning of the session, they received a short introduction to BabelDr and tested a few interactions. It was strongly suggested that they use complete sentences and ask yes/no questions, so that the patients could answer nonverbally.

Tool and Interface

Doctors only had access to the BabelDr tool. The diagnostic domain was set to “lower back pain” to match the patient complaint. In the context of this study, the other domains were not made available in order to simplify system usage. It was ascertained beforehand that all available questions potentially relevant to the patient complaint were included in this domain. The language pair was French to Arabic; the male or female patient was chosen depending on the case.

Data Collection and Analysis

Diagnoses

During the sessions, the doctors wrote down the information they were able to collect based on the patient’s responses. At the end of each session, the doctors wrote down their diagnoses. These data allowed us to answer the first question on whether the system enables doctors to reach a correct diagnosis.

System Usage

All interactions with the system were logged. For each session, we collected audio recordings of each spoken interaction with the system as well as the corresponding recognition results. We also logged which recognition results or text examples the doctors chose to translate for the patients. Finally, the duration of each session was measured. These data were analyzed to provide a quantitative answer to our second research question, namely, whether speech interaction is useful in this type of tool.

User Satisfaction

At the end of each session, participants completed a satisfaction questionnaire that included a total of 23 questions. The questions were derived from the System Usability Scale questionnaire by Brooke [18] and adapted to the functionalities of BabelDr, especially the speech and core sentence mapping aspects. Questions covered usability and learnability aspects of the BabelDr system during the study (7 items), appropriateness of the system to confidently reach a diagnosis (6 items), the speech component of the system (3 items), and the user's opinion regarding the usefulness of such a system in their daily medical practice (7 items). A 5-point Likert scale ("strongly disagree," "disagree," "neutral," "agree," and "strongly agree") was used to rate agreement with question items. These data contribute to a qualitative answer to our second research question.

Participants

Doctors

Study participants were 12 French-speaking doctors: 6 from the emergency service at HUG and 6 general practitioners who also regularly work in this service. All work in French, but three were not native speakers (#6, #11, #12). Only one doctor (#6) had previously used a former version of BabelDr in another study [5].

Standardized Patients

Of the two Arabic standardized patients, one was a man from Syria and one was a woman from Jordan. Both were refugees and recruited from among master's degree students at the Faculty of Translation and Interpreting. They had a high level of literacy, but no specific medical knowledge. Neither of the patients spoke French. One week before the experiment, both patients received an a priori list of symptoms for the condition they were to present, expressed in layman's terms. They were instructed to provide a negative or noncommittal answer to questions relating to other symptoms during the diagnostic interview.

All participants received remuneration for their participation in the study.

Ethical Considerations

The institutional ethics committee approved the study protocol (Req-2017-00996). Participation in the study was voluntary, with written agreement obtained from all doctors and patients. All data were anonymous and stored on a secure University of Geneva server.

Results

Diagnoses

Doctors were able to reach a correct diagnosis in all 24 sessions based on the information collected using BabelDr. For the renal colic scenario, four doctors proposed multiple related diagnoses (Table 1). These results showed that BabelDr was suitable for the task and allowed doctors to collect information reliably.

Textbox 1 gives examples of the most frequently asked questions for each scenario. In total, more questions were translated for the renal colic scenario than for the cystitis one (170 vs 126 unique interactions, respectively), probably reflecting the fact that the first scenario was more complex due to a larger number of possible related diagnoses and thus required more different questions.

Table 1. Diagnoses made by the 12 doctors.

Doctor no.	Female patient (with cystitis)		Male patient (with renal colic)	
	Diagnosis	Other diagnoses	Diagnosis	Other diagnoses
1	Cystitis	No	Renal colic	Pyelonephritis
2	Cystitis	No	Renal colic	No
3	Cystitis	No	Renal colic	No
4	Cystitis	No	Renal colic	No
5	Cystitis	No	Renal colic	No
6	Cystitis	No	Renal colic	Lumbosciatica
7	Cystitis	No	Renal colic	No
8	Cystitis	No	Renal colic	No
9	Cystitis	No	Renal colic	Pyelonephritis, lumbosciatica
10	Cystitis	No	Renal colic	No
11	Cystitis	No	Renal colic	No
12	Cystitis	No	Renal colic	Pyelonephritis, lumbosciatica, appendicitis

Textbox 1. Most frequently translated core sentences for each scenario, sorted by frequency.

Female patient with cystitis:

- *Pouvez-vous me montrer avec le doigt où est la douleur?* [Could you point with your finger to where it hurts?]
- *Avez-vous déjà eu ce type de douleur?* [Have you already had this type of pain?]
- *Bonjour* [Hello]
- *Je suis le docteur* [I’m the doctor]
- *Quand vous urinez, est-ce que ça brûle?* [Do you feel a burning sensation when you urinate?]
- *Avez-vous eu de la fièvre dernièrement?* [Have you had fever recently?]
- *Je vais m'occuper de vous aujourd'hui* [I will take care of you today]
- *Avez-vous mal au niveau des reins?* [Do you have pain in the kidney area?]
- *Je vais vous poser des questions avec cet appareil* [I will use this machine to ask you some questions]
- *Vos urines sont-elles rouges?* [Is your urine red?]
- *Êtes-vous d'accord?* [Do you agree?]
- *Il y a combien de semaines que vous avez eu vos dernières règles?* [How many weeks ago did you have your last period?]
- *Avez-vous été traité par antibiotique pour l'infection urinaire?* [Have you had antibiotic treatment for a urinary tract infection?]
- *Avez-vous eu une infection urinaire dernièrement?* [Have you recently had a urinary tract infection?]
- *Avez-vous des allergies connues?* [Do you have any known allergies?]

Male patient with renal colic:

- *Bonjour* [Hello]
- *Je suis le docteur* [I’m the doctor]
- *Vos urines sont-elles rouges* [Is your urine red?]
- *Avez-vous déjà eu ce type de douleur* [Have you had this kind of pain before?]
- *Pouvez-vous me montrer avec le doigt où est la douleur* [Could you point with your finger to where it hurts?]
- *Avez-vous eu de la fièvre dernièrement* [Have you had fever recently?]
- *Avez-vous mal au niveau des reins* [Do you have pain in the kidney area?]
- *La douleur aux reins irradie-t-elle vers un autre endroit* [Does the pain in the kidney area spread to any other place?]
- *Quand vous urinez, est-ce que ça brûle* [Do you feel a burning sensation when you urinate?]
- *Je vais vous poser des questions avec cet appareil* [I will use this machine to ask you some questions]
- *La douleur aux reins est-elle continue* [Is the pain in the kidney area continuous?]
- *Êtes-vous d'accord* [Do you agree?]
- *Je vais m'occuper de vous aujourd'hui* [I will take care of you today]
- *Depuis combien de jours avez-vous mal aux reins* [For how many days have you had pain in the kidney area?]
- *Avez-vous de la fièvre* [Do you have fever?]

Analysis of Interactions

For each doctor, we measured the time to complete the dialogue, the number of speech interactions, the number of speech interactions resulting in a translation for the patient, and the number of text items directly translated from the list of sentences. Table 2 shows that both the median time and the median number of translated speech interactions were higher for the renal colic scenario (16 min for 26 interactions) than for the cystitis scenario (13 min for 19 interactions), confirming the fact that the renal colic scenario was more complex.

Table 2 shows that doctors translated both speech and text, but used more speech interactions, suggesting that speech was generally preferred to text. The median number of speech interactions per dialogue that led to translations was 28.5 for the cystitis scenario and 36 for the renal colic scenario, whereas the median numbers for text interactions were 4.5 and 10, respectively.

Table 2. Time and number of interactions for both scenarios.

Variable	Female patient with cystitis, median (range)	Male patient with renal colic, median (range)
Time to diagnosis (min:seconds)	13:37 (4:09-35:37)	16:37 (4:35-23:35)
Speech interactions (n)	28.5 (17-46)	36 (20-66)
Speech translated (n)	19.5 (8-23)	26.5 (13-51)
Text translated (n)	4.5 (0-36)	10 (0-23)

Figures 3 and 4 present the interactions by participant and show that some used the text mode more often than others and that the number of speech sentences sent to translation differed from one participant to another. For different doctors, the proportions of recognition results leading to translations varied from 40% (8/20) to 94% (16/17) for the cystitis scenario and 37% (13/35) to 100% (20/20) for the renal colic scenario.

The association between the percentage of translated speech and the number of translated texts was investigated using a linear regression model. Since each medical practitioner assessed two patients, data were clustered. Therefore, a regression model with mixed effects was used: A random effect was set on the intercept to account for between-practitioner variability. In addition, a multivariable analysis was conducted to adjust for the session and the scenario.

Figure 3. Interactions by participant for the scenario with the female patient.

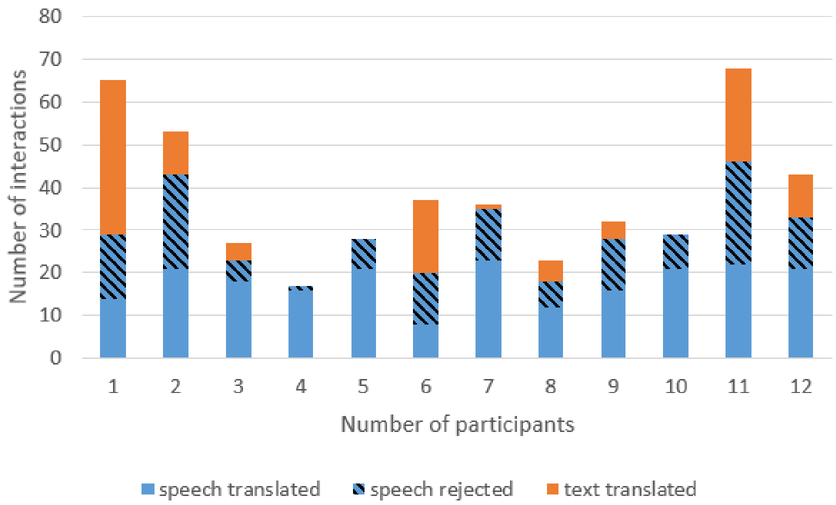
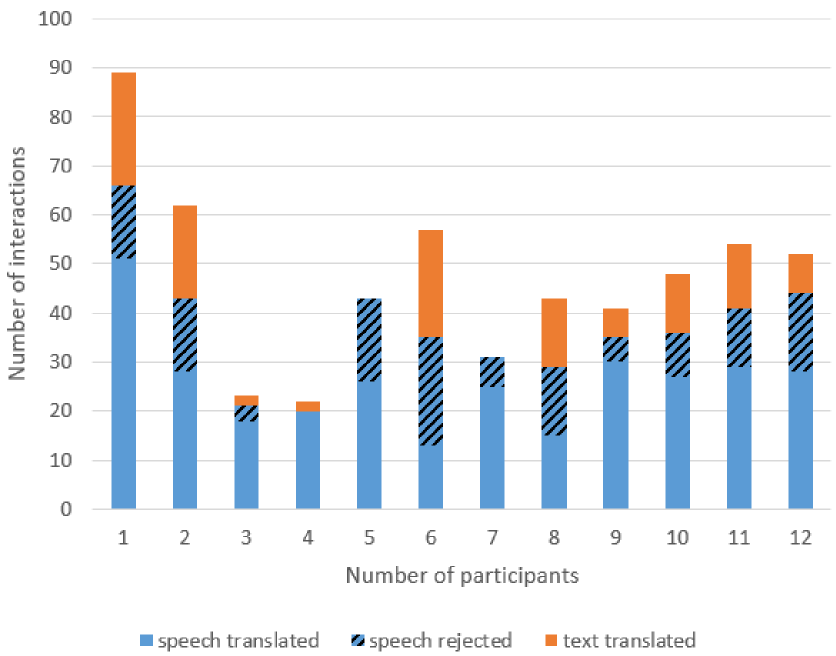


Figure 4. Interactions by participant for the scenario with the male patient.



The percentage of translated speech was negatively associated with the number of translated texts (*P*=.02): When the percentage of translated speech increased by 10%, the number of translated texts decreased by 2.6 (95% CI 0.7-4.4). After adjustment for the session and scenario, the decrease in the number of translated texts was similar (2.4; 95% CI 0.7-4.2; *P*=.02). This association is illustrated in Figure 5A. These results show that users who are not well recognized tend to use the text interface more often, thereby confirming the usefulness of including both modalities in such a tool.

The percentage of translated speech was higher in the second session than in the first session (difference=4.3%; 95% CI 1.1-7.4; *P*=.03). One possible interpretation may be that users familiarized themselves with system coverage in the first session and therefore used more coverage utterances in the second session, leading to better recognition of results and thus more translations.

Analyses by scenario showed that the proportion of translated speech was lower in the renal colic scenario than in the cystitis scenario (difference=4.3%; 95% CI –7.6 to –1.1; *P*=.03). This may be due to different factors such as concepts not covered by the system at the time of the study or errors in speech recognition or mapping to the core sentences (eg, cases where a sentence is badly recognized and therefore mapped to a different sentence). In some cases, the core sentence could also be too general or specific or considered inappropriate in the context. Table 3 presents some examples of these cases.

Figure 5. Association between the percentage of translated speech and the number of translated texts (A) and between French native speakers and the percentage of translated speech (B), system confidence score (C), and speech interaction (D). Circles represent each individual doctor's data; the black line represents the unadjusted regression line and black squares represent the mean values.

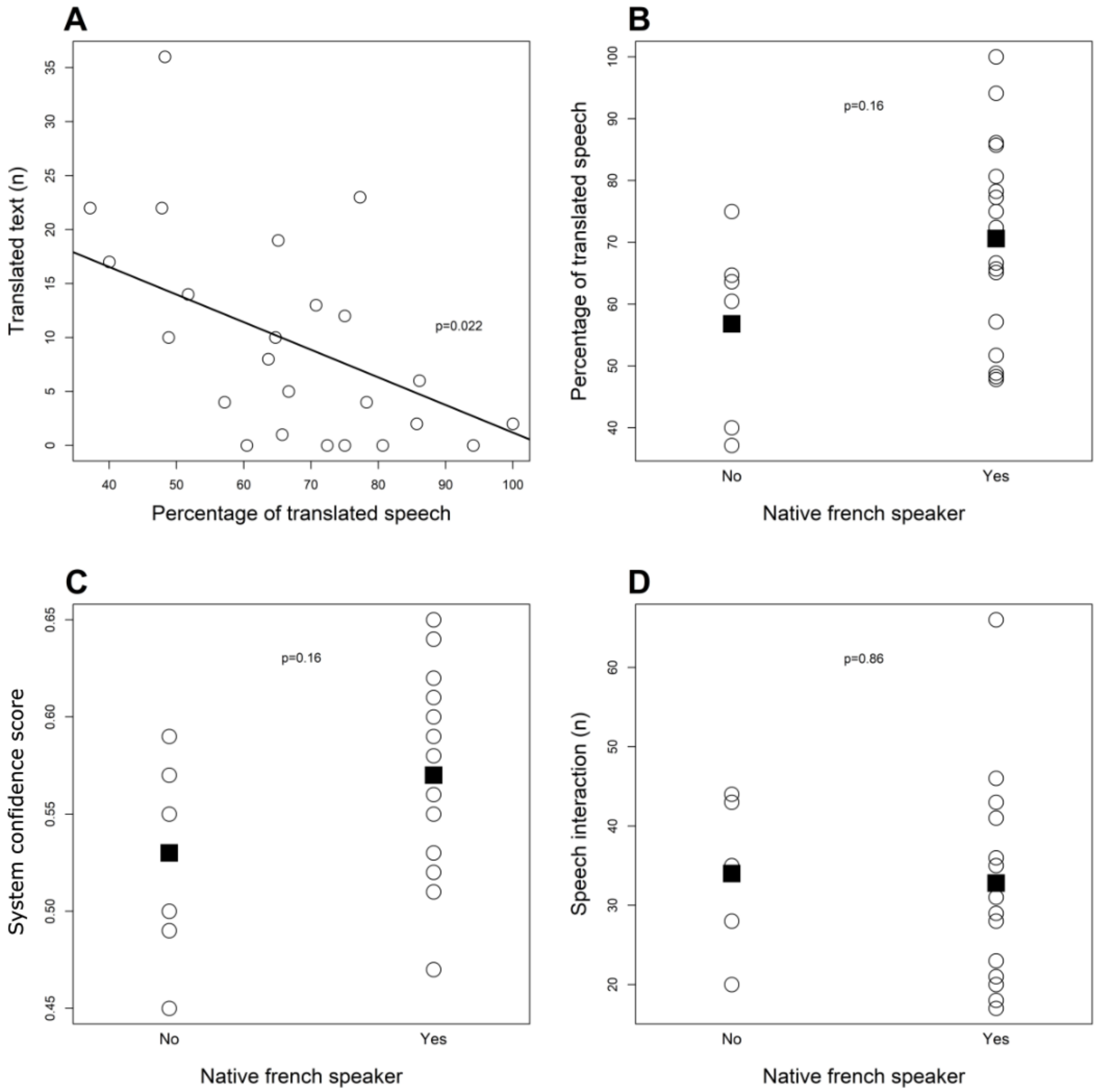


Table 3. Examples of transcriptions and mapped core sentences.

Speech utterances	Core sentences
Sent to translation	
Est-ce que vous avez fait du sport (Did you practice any sports?)	*Faites-vous de l'exercice* physique (Do you exercise?)
Est-ce qu'il y a du sang dans les urines (Is there any blood in your urine?)	*Les urines sont-elles rouges* (Is your urine red?)
Est-ce que vous avez beaucoup transpiré (Did you sweat a lot?)	*Suez-vous plus que d'habitude* (Do you sweat more than usual?)
Est-ce que c'est aujourd'hui (Is it today?)	*Avez-vous mal depuis aujourd'hui* (Do you have the pain since today?)
Est-ce que vous avez des pertes vaginales particulières (Have you observed any particular vaginal discharges?)	*Avez-vous des pertes blanches en dehors des règles* (Have you observed any white discharges outside normal menstruation?)
Not sent for translation by at least one doctor	
Avez-vous bu (Have you had anything to drink?)	*Avez-vous bu de* l'alcool (Have you consumed any alcohol?)
Est-ce que vous pourriez être enceinte (Could you be pregnant?)	*Êtes-vous enceinte* (Are you pregnant?)
Avez-vous du prurit (Do you have pruritus?)	*Avez-vous des* démangeaisons (Do you have itchiness?)
La douleur est-elle constante (Is the pain constant?)	*La douleur est-elle continue* (Is the pain continuous?)

Associations between French native speakers and the percentage of translated speech, system confidence, and speech interaction were also investigated using a linear regression model with fixed effects. No association was found between French native speakers and the percentage of translated speech (*P*=.16), system confidence (*P*=.16), and speech interaction (*P*=.86). Figure 5B-D illustrates these numbers. These results suggest that system performance is not significantly impaired by different accents.

User Satisfaction

Figures 6 and 7 show the results for seven questions related to the usefulness of the system for the diagnostic task and speech recognition included in the satisfaction questionnaires completed by the doctors after each dialogue (24 completed questionnaires). Overall, the doctors were satisfied with the speech interaction function and the usefulness of the system in the test context (19 negative, 54 neutral, and 116 positive judgments). All doctors considered that the system helped them reach a conclusion (Q3). They also liked the way the recognition result was presented (only one participant disagreed), which showed that they found the translation to the core sentence useful. All doctors thought that the system recognized their voice easily (Q4), and most believed that the system helped them to pose the question in a different way when the question could not be recognized (Q6: only 3 "disagree"). The most frequent criticism was that some doctors felt constrained by the tool (n=9/24) and were unable to ask all the questions they wanted to (5/24). In this respect, we observed differences between the two scenarios, suggesting that this issue is related to the system coverage or mapping of sentences to core sentences. Finally, all doctors believed that they could integrate such a system in their daily practice (Q7: no "disagree" or "strongly disagree").

Figure 6. Results of the satisfaction questionnaire completed after the dialogue with the female patient. The numbers in circles represent the number of doctors.

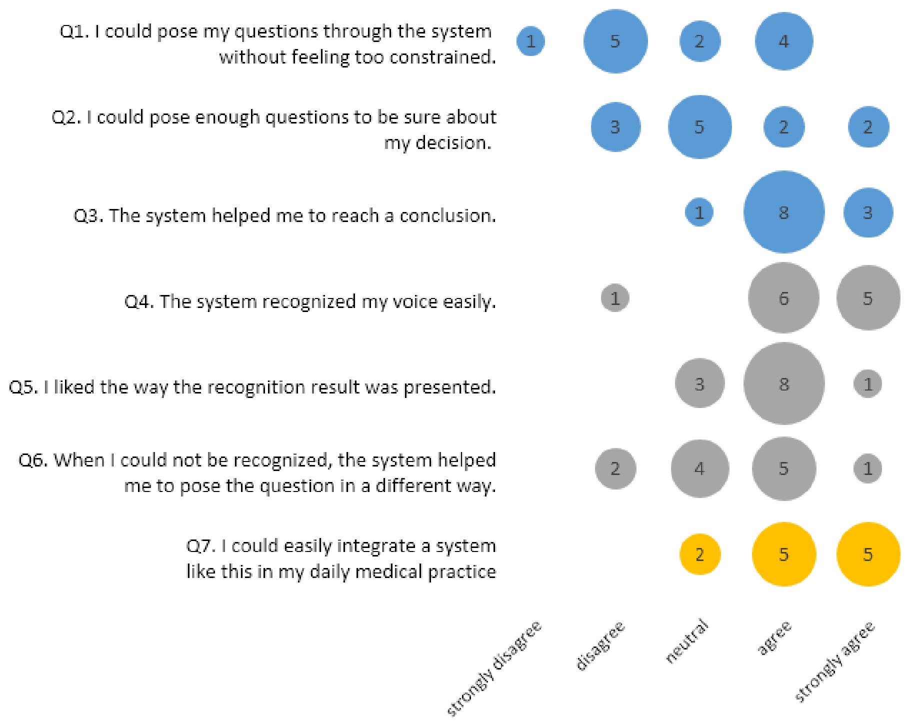
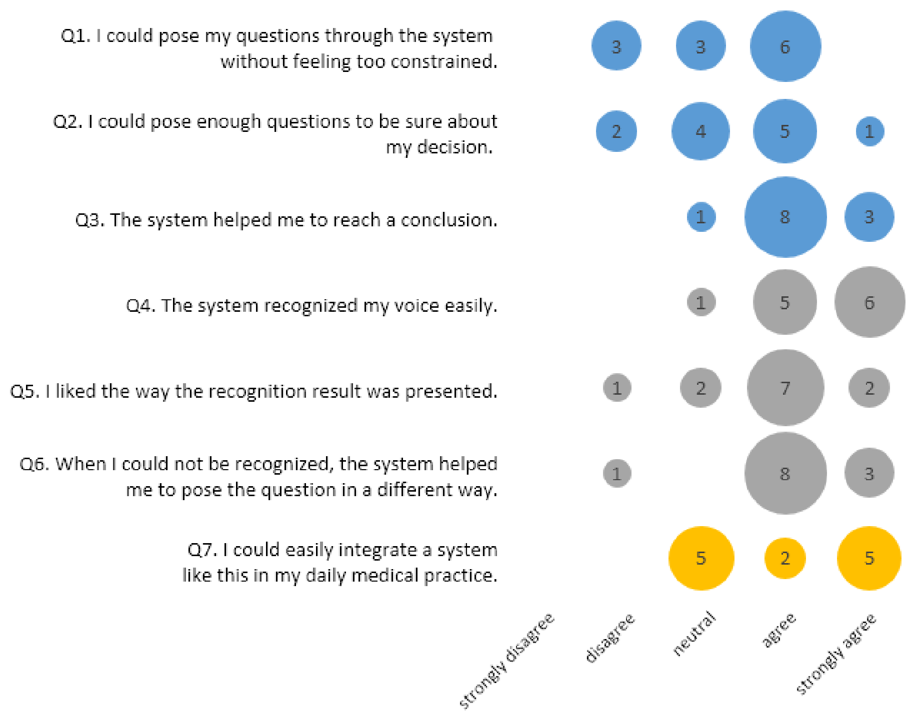


Figure 7. Results of the satisfaction questionnaire completed after the dialogue with the male patient. The numbers in circles represent the number of doctors.



Discussion

Principal Results

All participants were able to pose their questions to the patients and reach the correct diagnosis based on the information collected using BabelDr. However, although they believed that the system helped them to reach a conclusion, some felt constrained by the tool, as they could not ask enough questions to reach a diagnosis. Speech was the preferred modality, even if all doctors translated items from the text list, thus showing that both modalities are useful. The use of text was statistically influenced by the percentage of successful speech interactions and by the session (first use vs second use). Therefore, speech seems to help in using the system, as participants can express themselves freely and see the most related core sentences.

Comparison With Previous Research

Other studies have analyzed user satisfaction (of both patients and medical staff) [19,20] or the quality of translation with translation systems [4]. However, to our knowledge, this study is the first to measure the impact of the medium on diagnosis. This study confirmed the results of two previous evaluations of BabelDr. A comparison with a traditional fixed-phrase translator (Medibabble) in artificial settings (doctors had to find answers to specific questions) [21] showed that speech improves both usability (reduces time and number of clicks required to ask a question) and satisfaction. Another study [5,22] compared an earlier version of BabelDr with Google Translate at the level of diagnosis, satisfaction, and translation quality in a setting similar to this study. The main result was that BabelDr produced a better translation quality, improved precision (odds ratio: 0.04, 95% CI 0.02-0.12; $P<.001$ in favor of BabelDr) and fluidity (odds ratio: 0.04, 95% CI 0.02-0.10; $P<.001$ in favor of BabelDr) and led to more correct diagnoses than Google Translate.

Limitations

A preliminary version of the tool was used in the study. The system coverage, that is, the questions available to the doctors, is being continually improved based on the collected data. It is possible that the perception of constraint reported by the users was at least partially caused by insufficient coverage for the scenarios selected for this study, rather than by the system itself.

For the cystitis scenario, doctors would have benefited to have been able to change to another domain (abdominal pain), which was not accessible for this study. In addition, the doctors were informed beforehand of the patient's chief complaint. This matches the usual practice at HUG where this information is collected from patients during admission, but another study without prior information would ascertain whether the subdivision into domains, as done in BabelDr, meets the doctor's requirements.

The two standardized patients had a higher education level and no difficulty understanding the Arabic translations provided by the system. In the case of less literate patients, misunderstandings might cause incorrect patient responses and thus lead to incorrect diagnoses. Although the BabelDr translations are aimed at simplicity, a study of the translation quality and accessibility is currently in progress to ascertain whether the translations are suited to patients of different ages, education levels, and cultural and geographic origins.

Due to the rehearsed nature of the patient narratives, based on the given lists of symptoms rather than the potentially vague or contradicting observations by a real patient, it can be argued that the system performance in terms of diagnostic success would be lower with real patients. However, we suspect that the system's restriction to yes/no questions might actually improve clarity by enforcing precise questions and unambiguous patient responses.

During this experiment, we observed very few user errors, such as doctors forgetting to shut off the microphone or using questions that could not be answered nonverbally. Anecdotally, we have observed more such errors in real-use cases with real patients. However, it is possible that in the artificial setting of this study, doctors were more attentive to the system than when using it with a real patient, where the focus would be more on the patient, and thus, the proportion of successful interactions might be lower.

The number of dialogues per doctor (n=2) in this study was insufficient to measure a quantifiable learning effect, but a study is currently in progress at HUG, where BabelDr is used in real settings and the collected data will allow us to study its learnability.

Future Research

Our results show that speech and text interaction are complementary in a tool such as BabelDr. Future developments of the system include an improved text-search module providing more flexibility than the current keyword search.

Development of a bidirectional version of the system is ongoing. In this new version, patients will have an interface where they are presented with a range of responses (eg, numeric values, colors, and pictograms). This will allow us to extend the questions available to the doctors by including open questions and will possibly reduce doctors' feelings of being constrained by the system.

Conclusions

This study showed that a phraselator can be an alternative to machine translation and traditional fixed-phrase translators to reliably collect information from the patient in situations where no interpreter is available. Although doctors felt constrained by the system, they were able to confidently reach a diagnosis, and all believed they could use this type of system in everyday medical practice. The relevance of task-based evaluation to assess the usefulness and usability of translation tools for the diagnosis task was also demonstrated and confirms the importance of reliability in this type of oral context. Doctors clearly appreciated the way in which speech recognition results were presented in the form of a back translation to French, which provided the exact meaning of the translation produced for the patient. Future studies with BabelDr have to confirm these conclusions in real-life settings and investigate the proportion of cases that can be reliably diagnosed with such a tool.

Acknowledgments

The study was funded by the Fondation privée des Hôpitaux Universitaires de Genève. We thank Nuance Communication, Inc, for making their software available free of charge for research purposes. We also thank Jessica Rochat, Rosemary Sudan, and Emmanuel Rayner for their contributions to this study.

Conflicts of Interest

None declared.

References

1. Flores G, Laws MB, Mayo SJ, Zuckerman B, Abreu M, Medina L, et al. Errors in medical interpretation and their potential clinical consequences in pediatric encounters. Pediatrics 2003 Jan;111(1):6-14. [Medline: 12509547]
2. Rechel B, Mladovsky P, Devillé W, Rijks B, Petrova-Benedict R, McKee M. Migration And Health In The European Union (european Observatory On Health Systems And Policies). Maidenhead, Berkshire, England: Open University Press; 2019.
3. Wu AC, Leventhal JM, Ortiz J, Gonzalez EE, Forsyth B. The interpreter as cultural educator of residents: improving communication for Latino parents. Arch Pediatr Adolesc Med 2006 Nov;160(11):1145-1150. [doi: 10.1001/archpedi.160.11.1145] [Medline: 17088518]
4. Patil S, Davies P. Use of Google Translate in medical communication: evaluation of accuracy. BMJ 2014 Dec 15;349:g7392 [FREE Full text] [doi: 10.1136/bmj.g7392] [Medline: 25512386]
5. Bouillon P, Gerlach J, Spechbach H, Tsourakis N, Halimi S. BabelDr vs Google Translate: a user study at Geneva University Hospitals (HUG). In: Proceedings of the 20th Annual Conference of the European Association for Machine Translation. 2017 Presented at: 20th Annual Conference of the European Association for Machine Translation; 29-31 May 2017; Prague, Czech Republic p. 747-752.
6. Dew K, Turner A, Choi Y, Bosold A, Kirchhoff K. Development of machine translation technology for assisting health communication: A systematic review. J Biomed Inform 2018 Sep;85:56-67. [doi: 10.1016/j.jbi.2018.07.018] [Medline: 30031857]
7. BabelDr: Spoken Language Translation of Dialogues in the Medical Domain. URL: https://babeldr.unige.ch/ [accessed 2019-04-23] [WebCite Cache ID 74jFPa3QF]
8. Bouillon P, Flores G, Starlander M, Chatzihrisafis N, Santaholma M, Tsourakis N, et al. A bidirectional grammar-based medical speech translator. In: Proceedings of the Workshop on Grammar-Based Approaches to Spoken Language Processing. 2007 Presented at: ACL 2007; June 2007; Prague, Czech Republic p. 41-48.
9. Wikipedia. Phraselator URL: https://en.wikipedia.org/wiki/Phraselator [accessed 2019-04-23] [WebCite Cache ID 76RzhshCe]
10. Ehsani F, Kimzey J, Zuber E, Master D, Sudre K. Speech to speech translation for nurse patient interaction. In: Proceedings of the workshop on Speech Processing for Safety Critical Translation and Pervasive Applications. 2008 Presented at: 22nd International Conference on Computational Linguistics (Coling 2008); August 2008; Manchester, UK p. 54-59.
11. Medibabble. URL: http://medibabble.com/ [accessed 2019-04-23] [WebCite Cache ID 74jFaAVTG]
12. Universal Doctor. URL: http://www.universaldoctor.com/ [accessed 2019-04-23] [WebCite Cache ID 74jFhoKAS]
13. Rayner M, Armando A, Bouillon P, Ebling S, Gerlach J, Halimi S, et al. Helping domain experts build phrasal speech translation systems. In: Quesada JF, Martín Mateos FJ, Lopez-Soto T, editors. Future and Emergent Trends in Language Technology. First International Workshop, FETLT 2015, Seville, Spain, November 19-20, 2015, Revised Selected Papers. Cham, Switzerland: Springer International Publishing; 2016.
14. Gerlach J, Spechbach H, Bouillon P. Creating an online translation platform to build target language resources for a medical phraselator. In: Proceedings of the 40th Edition of the Translating the Computer Conference. 2018 Presented at: 40th Edition of the Translatingthe Computer Conference (TC 40); November 2018; London, UK p. 60-65.
15. Rayner M, Tsourakis N, Gerlach J. Lightweight spoken utterance classification with CFG, tf-idf and dynamic programming. In: Statistical Language and Speech Processing. Cham, Switzerland: Springer International Publishing; 2017:143-154.
16. Seligman M, Dillinger M. Automatic speech translation for healthcare: some internet and interface aspects. In: Proceedings of the 10th International Conference on Terminology and Artificial Intelligence (TIA-13). 2013 Presented at: 10th International Conference on Terminology and Artificial Intelligence (TIA-13); October 28-30; Paris, France.
17. Gao Y, Gu L, Zhou B, Sarikaya R, Afify M, Kuo K, et al. IBM MASTOR SYSTEM: Multilingual automatic speech-to-speech translator. In: Proceedings of the First International Workshop on Medical Speech Translation. 2006 Presented at: NAACL/HLT; June 4-9; New York, NY, USA.
18. Brooke J. SUS: a quickdirty usability scale. In: Jordan PW, Thomas B, Weerdmeester BA, McClelland AL, editors. Usability evaluation in industry. London: Taylor and Francis; 1996:189-194.
19. Albrecht U, Behrends M, Schmeer R, Matthies H, von Jan U. Usage of multilingual mobile translation applications in clinical settings. JMIR Mhealth Uhealth 2013 Apr 23;1(1):e4 [FREE Full text] [doi: 10.2196/mhealth.2268] [Medline: 25100677]
20. Taicher B, Alam R, Berman J, Epstein R. Design, implementation, and evaluation of a computerized system to communicate with patients with limited native language proficiency in the perioperative period. Anesth Analg 2011 Jan;112(1):106-112. [doi: 10.1213/ANE.0b013e3182009cd1] [Medline: 21081767]
21. Boujon V, Bouillon P, Spechbach H, Gerlach J, Strasly I. Can speech-enabled phraselators improve healthcare accessibility? A case study comparing BabelDr with MediBabble for anamnesis in emergency settings. In: Proceedings of the 1st Swiss Conference on Barrier-free Communication. 2017 Presented at: 1st Swiss Conference on Barrier-free Communication; September 2017; Winterthur, Switzerland p. 32-38. [doi: 10.21256/zhaw-3000]

22. Spechbach H, Halimi S, Gerlach J, Tsourakis N, Bouillon P. Comparison of the quality of two speech translators in emergency settings: A case study with standardized Arabic speaking patients with abdominal pain. In: Proceedings of the 11th European Congress of Emergency Medicine. 2017 Presented at: 11th European Congress of Emergency Medicine; September; Athens, Greece.

Abbreviations

HUG: Geneva University Hospitals

Edited by C Lovis; submitted 17.12.18; peer-reviewed by M Seligman, P Boisrond, K Day, A Cox, J Liang; comments to author 12.01.19; revised version received 29.03.19; accepted 15.04.19; published 07.05.19

Please cite as:
Spechbach H, Gerlach J, Mazouri Karker S, Tsourakis N, Combescure C, Bouillon P
A Speech-Enabled Fixed-Phrase Translator for Emergency Settings: Crossover Study
JMIR Med Inform 2019;7(2):e13167
URL: http://medinform.jmir.org/2019/2/e13167/
doi: 10.2196/13167
PMID: 31066702

3.3.2 Fourth article: "User satisfaction with a speech-enabled translator in emergency settings" (4)

We conducted a real-life test (phase II cohort) of BabelDr tool in the ED with random patients who accepted to participate, had any language in common with the caregivers, and spoke a language included in BabelDr. At the beginning, all patients received information in their native language about the study, the BabelDr tool, and the fact that they could stop the experiment at any moment. The doctors could use as they preferred the voice recognition and the text to translate the sentences. All medical doctors and patients received an introduction about the tool. We included 22 patients with the following languages: Spanish (9); Farsi (6); Arabic (4); Tigrinya (2); Albanian (1). Patient diagnosis was varied and ranged from digestive to respiratory issues, cutaneous infection, urinary infection and articular problems. At the end of the test, both patients and medical doctors competed a survey about their satisfaction with the tool in their native language. We received positive feedback of 85% from patients and medical doctors. Doctors estimated that BabelDr allowed them to understand the patient in 90.1% of cases, and 68.2% of patients answered that BabelDr allowed them to completely explain the reason for the consultation. The association between doctors' satisfaction and the tool was correlated with the ability to be able to interact orally with the system (p=0.005). Approximately one-half of patients would have preferred a human interpreter. Some medical doctors described being constrained by the system and misunderstood translations by patients, mainly related to low health literacy.

Regarding the fact that 53.3% of patients expressed that they would have preferred an interpreter, we observed similar results in another study with a cross-over design conducted at HUG (12) where we included 8 French-speaking medical doctors and 8 English-speaking standardized patients based on two scenarios (appendicitis and cholescystitis). In the satisfaction survey, we observed that the overall satisfaction rate was higher with telephone interpreters, except for confidentiality. This highlighted the fact that BabelDr is satisfactory for use in the ED for specific domains, such as triage and a confidential interview, and complementary with other mediums.

Digital Personalized Health and Medicine
L.B. Pape-Haugaard et al. (Eds.)

doi:10.3233/SHTI200472

User Satisfaction with a Speech-Enabled Translator in Emergency Settings

Antony Akash JANAKIRAM[a], Johanna GERLACH[b], Alyssa VUADENS-LEHMANN[a], Pierrette BOUILLON[b] and Hervé SPECHBACH[a]
[a] *Geneva University Hospitals, Department of Community Medicine, Geneva, Switzerland*
[b] *Faculty of translation and interpreting (FTI), University of Geneva, Geneva, Switzerland*

Abstract. In medical emergency situations, the language barrier is often a problem for healthcare quality. To face this situation, we developed BabelDr, an innovative and reliable fixed phrase speech-enabled translator specialised for medical language. Majority of participants (>85%) showed a positive satisfaction level using BabelDr.

Keywords. Fixed phrase speech-enabled translator, web application, user satisfaction, emergency settings

1. Introduction

Following the refugee crisis, European hospitals are increasingly confronted with patients who have no language in common with the medical staff. At the Geneva University Hospitals (HUG), 52% of the patients are foreigners and 8% speak no French at all. Miscommunication is a known threat to quality, security and equitability of medical care [1, 5]. Professional interpreters are not always available in emergency settings and considered to be expensive [2]. Generic technologies such as Google Translate (GT) are not reliable enough [3]. Medical fixed phrase translator lack of usability. In this context, the HUG in collaboration with the FTI, supported by the Private foundation of HUG, developed BabelDr, a speech-enabled fixed-phrase translator. Like other systems, it relies on pre-translated sentences, but includes speech recognition, allowing doctors to freely ask questions instead of searching for them in a list. Pretest showed that BabelDr is significantly more precise than GT and presents higher usability than MediBabble [4]. The aim of the present study is to evaluate BabelDr in real emergency settings. We focus on two aspects: global user satisfaction and the perceived quality of patient-doctor communication through the system, from the patient's and doctor's point of view.

2. Method

Patients were recruited between January and August 2019 in the outpatient emergency unit of the HUG. Inclusion criteria are: patients must speak a language available on BabelDr and must not have a French level sufficient for a standard diagnostic interview.

Patients are recruited in the outpatient emergency unit of the HUG. Exclusion criteria are patients who aren't able to read in their language. Each doctor followed a short training before using BabelDr. Participants are enrolled during daytime, fill a consent form and receive instructions on how to interact with the tool. During the consultation, the use of the tool can be interrupted by the doctor or the patient. Doctors can use the speech recognition system or directly select sentences in a list to ask their questions and the patients are required to answer non-verbally. After the consultation, all participants (doctors and patients) fill in a satisfaction survey paragraph.

3. Results

In total, 22 patients (age: *M*=41.18, *SD*=17.44) are eligible for this study and speak: Spanish (9), Arabic (4), Farsi (6), Tigrinya (2) and Albanian (1). Results on the satisfaction level show predominantly a positive feedback from both patients and doctors (>85%). Regarding patient-doctor communication, 68.2% of patients thinks the system allows them to completely explain the reason for consulting and 90.1% of doctors think BabelDr allows them to understand the patient's problem.

4. Discussion

We observe that most patients are able to communicate their reason to visit and doctors mostly understand the problem, suggesting that BabelDr is suitable for communication in emergency settings. User satisfaction is higher for patients than doctors, who can feel constrained by the available questions. Patients' dissatisfaction is related to misunderstood translations (due to level of literacy, text-to-speech quality and dialects).

5. Conclusion

Based on the remarks of patients and doctors, we continue to develop the tool by expanding coverage and adding more languages. A bidirectional system, allowing patients to answer by selecting pictograms, is also in development.

References

[1] G. Flores, M.B. Laws, S.J. Mayo, B. Zuckerman, M. Abreu, L. Medina, & E.J. Hardt, Errors in medical interpretation and their potential clinical consequences in pediatric encounters, *Pediatrics* **111**(1) (2003), 6-14.

[2] G. Major, J. Napier, L. Ferrara, & T. Johnston, Exploring lexical gaps in Australian Sign Language for the purpose of health communication, *Communication & Medicine* **9**(1) (2012), 37-47.

[3] S. Patil, & P. Davies, Use of Google Translate in medical communication: evaluation of accuracy, *Bmj* **349**, g7392

[4] H. Spechbach, J. Gerlach, S.M. Karker, N. Tsourakis, C. Combescure, & P. Bouillon, A Speech-Enabled Fixed-Phrase Translator for Emergency Settings: Crossover Study. *JMIR medical informatics* **7**(2) (2019), e13167

[5] M. Wasserman, M.R. Renfrew, A.R. Green, L. Lopez, A. Tan-McGrory, C. Brach, & J.R. Betancourt, Identifying and preventing medical errors in patients with limited English proficiency: key findings and tools for the field. *Journal for Healthcare Quality* **36**(3) (2014), 5-16.

4. Conclusion and perspectives

4.1 Conclusion

In the ED, patient arrival is not scheduled and non-urgent cases are at risk of a long wait duration, according to their classification on the Swiss Emergency Triage Scale. As described in this thesis, we are faced here with the paradox of patients suffering from minor disease who are even more sensitive to the quality of the wait environment. This problematic of managing non-urgent cases in a stressful environment is a source of frustration and aggressivity for both patients and caregivers.

This thesis describes instruments acting on the overall satisfaction, as well as the wait perception during throughput, and demonstrates that the wait perception directly influences overall satisfaction and confirms the well-described disconfirmation paradigm as a key point for satisfaction in the ED. Indeed, if the delta between the actual waiting time and the expected waiting time were >1, the perceived waiting time was estimated as "very long" or "long" and had a direct negative impact on the overall satisfaction (negative disconfirmation). Here, we show that it is possible to act on ED throughput with instruments permitting the patient to alter their perception of the waiting time.

Consistent with the literature, the two strongest predictors of the perceived waiting time before the medical consultation are the appropriate assessment of the emergency level by caregivers and the feeling of being forgotten. We need to analyze the psychological profile of these patients to better understand the reasons of feeling under-triaged and perhaps allow them to be more active on this aspect by permitting them to make the triage or a part of it by themselves for example.

Another important finding described is that when the wait is moved outside the ED, the duration of time does not influence the wait perception and satisfaction. When the wait perception is changed and is outside the ED, the duration of stay before seeing a doctor does not influence the satisfaction level. Of note, we observed that 28% of patients felt stressed by waiting outside.

We have discussed here the different mediums to overcome the language/communication barriers in the context of needs for an augmentative and alternative way of communication in the medical setting in the ED. The instruments developed show the need of natural language with voice recognition, the still-present risk of errors in the medical field with mediums like Google Translate, and the fact that current technology cannot yet replace the human for sophisticated discussion. However, given the current rapid progress of artificial intelligence (AI) possibilities, it is certain that MT will permit to discuss with patients quite naturally in the near future with a performant voice recognition. In this field, we need to have scientifically validated recommendations on how to use this future technology and strong comparison of different systems.

4.2 Perspectives

The four articles mentioned above illustrate several aspects of the development of instruments to address the patient's perception of waiting time to optimize and customize care provided, and the personal contribution of the author of this thesis. Two of the four articles emphasize the importance of using technology to enhance care for patients who do not share a common language with the caregiver. These studies represent four intertwined areas of research based on the three stages of the ED flow, i.e., input, throughput and output, as well as the issue of direct qualitative communication to address challenges of the language barrier, which has a direct impact on quality and satisfaction with care. These are all part of continuing ongoing research discussed below.

4.2.1 Input

The major limitation is that caregivers are often victims of the input, which is difficult to control (13) and often with large variations in patient influx, thus impacting on the quality of welcome, duration of stay, and overall satisfaction.

In the HUG outpatient emergency unit, we conducted an extensive analysis of input retrospective data and were able to observe a regular irregular pattern. For example, there were 20% more patients every Monday compared to other weekdays, with a daily peak of activity as described in other EDs (14). These data allowed to adapt and refine our staffing rotation, but this is not sufficient and AI may be able to help further in this area (15). For example, we are currently working on an AI project to predict the daily number of patients per hour 24h/24h with an anticipation of a few days (16). Using retrospective and anticipated data, we also plan to predict trends and patterns, such as the meteorological forecast (17), the number of X-rays, computed tomography scans, laboratory requirements, and the need for consultants per hour, also with a few days anticipation. Such an anticipation would help to adapt our internal capacity and should make it possible to adapt the working hours of care staff and better inform our partners, such as radiology, the laboratory, and various specialized consultations. The objective here is to be able to welcome the patients under good conditions, respecting their comfort and privacy, as well as acting on the perception of care and overall satisfaction. Such anticipation should also have a beneficial effect on quality of care and job satisfaction of caregivers (18).

In the field of patient engagement, Informed (https://www.hug.ch/application-mobile/infomed) (11), an app developed at HUG with the support of the HUG private foundation and the successor of Infokids (11), will permit the patient to document their symptoms, with an algorithm orienting them to the ED, a pharmacy, or a general practitioner. If they decide to come to the ED, they can already provide administrative information and be included in a virtual waiting room. The ultimate goal is to geolocalize patients, propose the shortest way to an ED considering the waiting time, time to travel, age and symptom characteristics. The motto of the app is "the right patient at the right time in the right place and customizing the care at each touch point" (19), with a focus on improving overall satisfaction, based on the principle that satisfaction has an impact on quality of care and adherence to treatment, as well as health and safety of care in general from a cost-efficiency perspective.

4.2.2 Throughput

Another perspective regarding the perception of waiting time in the context of the project of AI described above is to give not only information on the predicted waiting time to each patient and family arriving in the ED, but also information on the potential overall duration of stay, need for laboratory/radiological examinations, and the possibility if they will return home or not.

When patients wait outside the ED, there is a risk that they may feel forgotten or worry that their condition could worsen. In our study, 28% of patients felt stressful by waiting outside (2). The solution could be to allow the patient to contact a caregiver remotely and to be informed about their position in the queue, permitting him/her to chat with a caregiver (20) or a chatbot, and/or to have access to a device to regularly take vital signs and send these to the caregivers.

A plan with Informed is to inform patients during their stay in the ED about the prescription of laboratory examinations, thus offering the possibility to ask to wait outside the ED, while allowing to interact with caregivers. To ensure that Infomed remains a viable app for patients, we need to incorporate AI into the system in a second version.

4.2.3 Output

To impact on output, we plan to predict the need for a bed and to obtain an anticipation in this area as it has been reported that output is one of the most relevant elements of overcrowding. We also aim to predict an individualized trajectory in the ED for patients comprising a prediction of time spent in the ED, together with a prediction of complementary examinations, potentially leading to the need for hospitalization. It is indispensable to change the perception of care for patients and their families and to seek innovative ways to enhance overall patient satisfaction. With Infomed, we plan give the patient information at home, including the main diagnosis and different instructions, such as warning signs that should prompt a new consultation.

4.2.4 Different mediums to overcome language/communication barriers

Although there is an extensive body of information (21) in the literature on patient safety, patient satisfaction, the advantages/disadvantages of human interpreters, ad hoc interpreters, MT or phraselator, no study has analyzed the complementarity aspects of these different solutions, with the availability of a handout on how to use each one. Notably, the concept of literacy in MT is quite new and was introduced by Bowker and Buitrago Ciro in the context of writing an academic text. Training in this field is important for a reasoning utilization of different tools. The perspective will be to test different mediums and propose a “village” of complementary solutions (22). De Graaf and al highlight this point in a survey conducted in a Swiss French-speaking ED as 63% of doctors responded that they do not have any guidelines to communicate with allophone patients (presented in spring congress of SGAIM, Basel, 2024). In addition, most feel that they are treating allophone patients worse than French speakers. (This latter point will be addressed in more detail and analyzed in the thesis of Céline De Graaf, which I co-conduct with Prof. Pierrette Bouillon.)

The field of augmentative and alternative communication offers considerable opportunities for innovation and I plan to pursue several different axes of development. First, BabelDr and PictoDr will be merged, allowing to use BabelDr in many different domains. BabelDr with its 10,000 medical sentences related to more one million variations will be matched to the USMLS concept and then translated into pictographs. The objective is to use PictoDr for languages not covered by BabelDr and/or with patients needing an augmentative and alternative way of communication, such as those with a cognitive disorder or low health literacy. In the field of triage, BabelDr-PictoDr must be suitable and reliable with closed sentences. This needs further analysis and we are currently working on a research project in this area.

Importantly, we need to test the pictographs with different populations of patients (children, adults, seniors) from different cultural backgrounds and with caregivers from different experiences as we observed a very subjective interpretation. The challenge here is to find medical pictographs that can be understood in a cross-cultural setting. Another challenge is to develop pictographs understood by caregivers who have different backgrounds and/or cultural past history. A second challenge, as pictographs for medical communication are sparse, is to create homemade ones, which necessitates massive resources. A first study to evaluate the pictographs in BabelDr reported (5) that 983/3655 pictographs were found to be useful representations of the UMLS concept by at least two of the three medical students acting as evaluators. Pictographs representing physical objects such as medical devices were the most successful. This highlights the need for additional testing among community members and to consider offering the possibility to have several pictographs for the same UMLS concept, depending on the cultural context. One of the limitations here is the lack of pictographs concerning the medical domain. The next step on-going is to have pictographs generated by IA (IA generative).

The development of French sign language to the medical context is also ongoing and very innovative. In addition to the videos made one-by-one by a nurse familiar with sign language in French-speaking Switzerland, we are presently training an avatar. According to the Swiss Federal Statistical Office, there are approximately 1000 deaf people, including 1000 in Geneva. Although there are only around 10 such patients each year in our ED, this translates as 10 times where we encounter an unsatisfactory communication, with known consequences in terms of safety of care and overall satisfaction. Indeed, it is described (23) that deaf people encounter barriers when they need to communicate in a healthcare context and the risk of a misunderstood diagnosis and treatment between those patients and caregivers is obviously high. We need to develop communication tools as we lack professional human interpreters in this field.

However, the human videos require a major logistic investment, are time-consuming and costly to produce, and cannot be modified when finalized. The avatar (Figure 5) is cheaper, requires few logistics and we can modify the output at any time according to our needs. A clear disadvantage is that the avatar does not express emotions like a human and therefore we have lost facial expressions and eye tracking. For this reason, we are testing the human and the avatar with standardized deaf patients and a medical doctor by asking both of them how they can interact, what is their satisfaction level, and their preferred solution in the ED. We already have preliminary results showing that deaf people may prefer an avatar as they can choose if it is a man, a woman or of white or black ethnicity for example. In addition, as the community is not large, they would appreciate an avatar for confidential medical information. In a recent congress presentation (Handicap 2022, Paris), we demonstrated that the human videos are more comprehensible than the avatar. Another ongoing study in the ED context is to test the diagnostic performance using a human who signs and an avatar. We plan to extend this test in settings other than the ED, such as the HUG pharmacy, which is allowed to conduct consultations.

The human videos serve as a reference corpus with annotations concerning the motion, the animation of hands and facial expression permitting to generate and make a control of videos made with an avatar. We also use different sign database regarding details such the gaze (neutral, right, down) and the position of the head and shoulders, which permits to animate the avatar as close as possible to human videos. When there is a risk of misunderstanding, we cut the video into paraphrases and/or add subtitles. This is important as many medical concepts do not exist in sign language. Based on this we develop an online corpus for the sign language on Yareta, a rare one medical corpus for the sign language (https://yareta.unige.ch/archives/e93920a5-e5b8-47de-9979-d1fc594c068d).

The objective is to have a system with a mix of human videos and an avatar and one of the next steps would be to compare the satisfaction level of caregivers and deaf patients between the human videos and the avatar. The patients can answer with the pictographs developed in BabelDr (Figure 6).

Figure 6: The avatar of BabelDr, permitting the deaf patient to answer with the pictographs. (Reproduced from www.babeldr.unige.ch.)

Finally, the University of Geneva Commission for Research and Development has recently accepted a new project, which will help to further develop BabelDr and test for example Google Translate or other MT tools. The goal is to compile a corpus of interaction of the thousands of recorded videos from the learning stations of the University of Geneva Faculty of Medicine. The videos comprise evaluated interactions between standardized patients and medical students during simulated consultations for a specific reason. The project is to make a corpus by transforming the speech into text by annotating the year of the medical student, his/her gender, and the dysfluency, and to separate the speech into simple sentences with one meaning. This corpus will be hosted on an open platform (Yareta; https/yareta.unige.ch) with a digital object identifier permitting multiple axes of research and could also be suitable to develop a chatbot for students, permitting them to train before examinations. Notably, it will also be an important step forward to help improve communication efficiency in the ED, offer an enhanced quality of care for our patients, and improve the number of sentences recognized in BabelDr-PictoDr. This project is ongoing and has been recently accepted by the University of Geneva Commission for Ethical Research (CUREG).

5. References

1. Spechbach H, Rochat J, Gaspoz JM, Lovis C, Ehrler F. Patients' time perception in the waiting room of an ambulatory emergency unit: a cross-sectional study. BMC Emerg Med. 2019;19(1):41.
2. Ehrler F, Rochat J, Siebert JN, Guessous I, Lovis C, Spechbach H. Use of a semiautomatic text message system to improve satisfaction with wait time in the adult emergency department: cross-sectional survey study. JMIR Med Inform. 2022;10(9):e34488.
3. Spechbach H, Gerlach J, Mazouri Karker S, Tsourakis N, Combescure C, Bouillon P. A speech-enabled fixed-phrase translator for emergency settings: crossover study. JMIR Med Inform. 2019;7(2):e13167.
4. Janakiram AA, Gerlach J, Vuadens-Lehmann A, Bouillon P, Spechbach H. User satisfaction with a speech-enabled translator in emergency settings. Stud Health Technol Inform. 2020;270:1421-2.
5. Gerlach J, Bouillon P, Norré M, Spechbach H. Translating medical dialogues into pictographs: an approach using UMLS. Stud Health Technol Inform. 2023;302:823-4.
6. Taylor C, Benger JR. Patient satisfaction in emergency medicine. Emerg Med J. 2004;21(5):528-32.
7. Thompson DA, Yarnold PR. Relating patient satisfaction to waiting time perceptions and expectations: the disconfirmation paradigm. Acad Emerg Med. 1995;2(12):1057-62.
8. Cohen EL, Wilkin HA, Tannebaum M, Plew MS, Haley LL, Jr. When patients are impatient: the communication strategies utilized by emergency department employees to manage patients frustrated by wait times. Health Commun. 2013;28(3):275-85.
9. Shaikh SB, Jerrard DA, Witting MD, Winters ME, Brodeur MN. How long are patients willing to wait in the emergency department before leaving without being seen? West J Emerg Med. 2012;13(6):463-7.
10. Hedges JR, Trout A, Magnusson AR. Satisfied patients exiting the emergency department (SPEED) study. Acad Emerg Med. 2002;9(1):15-21.
11. Ehrler F, Lovis C, Rochat J, Schneider F, Gervaix A, Galetto-Lacour A, et al. [InfoKids: changing the patients' journey paradigm in an Emergency Department]. Rev Med Suisse. 2018;14(617):1538-42.
12. De Graaf C, Bouillon P, Spechbach H. Communicating in emergency settings: BabelDr or telephone interpreting? Stud Health Technol Inform. 2023;302:813-4.
13. He J, Hou XY, Toloo S, Patrick JR, Fitz Gerald G. Demand for hospital emergency departments: a conceptual understanding. World J Emerg Med. 2011;2(4):253-61.
14. Kang SW, Park HS. Emergency department visit volume variability. Clin Exp Emerg Med. 2015;2(3):150-4.
15. Harrou F, Dairi A, Kadri F, Sun Y. Forecasting emergency department overcrowding: A deep learning framework. Chaos Solitons Fractals. 2020;139:110247.
16. Calyps Saniia. Calyps SA. Data solutions. Sion, Switzerland (https://www.calyps.ch/fr/solutions-medicales-calyps/solution-calyps-saniia/).
17. Sudarshan VK, Brabrand M, Range TM, Wiil UK. Performance evaluation of Emergency Department patient arrivals forecasting models by including meteorological and calendar information: A comparative study. Comput Biol Med. 2021;135:104541.
18. Alexandrescu R, Bottle A, Jarman B, Aylin P. Classifying hospitals as mortality outliers: logistic versus hierarchical logistic models. J Med Syst. 2014;38(5):29.
19. Gierlinger S, Almulla N, Lee C, Lee T. What really matters to patients. NEJM Catalyst. 2024;5(4). doi:10.1056/CAT.24.0328.
20. Brennan I, Datta M, Miner A, Dibble B. Enhancing emergency department patient experience and flow through real-time text-messaging updates. NEJM Catalyst. 2024;5(4). doi:10.1056/CAT.23.0252.
21. Hudelson P, Chappuis F. Using voice-to-voice machine translation to overcome language barriers in clinical communication: an exploratory study. J Gen Intern Med. 2024;39(7):1095-102.
22. Hsieh E. Not just "getting by": factors influencing providers' choice of interpreters. J Gen Intern Med. 2015;30(1):75-82.
23. Scheier DB. Barriers to health care for people with hearing loss: a review of the literature. J N Y State Nurses Assoc. 2009;40(1):4-10.

6. Acknowledgements

I wish to express my heartfelt gratitude to Prof. Jean-Michel Gaspoz for his unwavering support, and to my current department head, Prof. Idris Guessous, whose inspiration, guidance, and vision have been invaluable. I extend my sincere thanks to my close colleagues, Dr med Chloé Chevallier-Lugon and Dr med Julien Salamun, who have been actively contributing to my research.

I extend my deepest gratitude to Prof. Pierrette Bouillon and her research team, with whom I have been collaborating weekly for many years on the subject of augmentative and alternative communication.

A special thanks goes to my beloved family—my wife, Sarah Rumer, and my son, Lionel—whose love and encouragement have been my constant source of strength. Special thanks also to my parents who have supported me since birth. I am also deeply grateful to those who have contributed to my research in various capacities, including Mrs Rosemary Sudan and Mrs. Michelle Le Braz.

I would like to acknowledge the Geneva Emergency Network and the dedicated medical and nursing teams for their exceptional support.

Finally, I would also like to extend my heartfelt thanks to Prof. Frauke Mueller for her supervision.