Health Informatics Journal

Volume 24 Number 2 June 2018

Article

Health Informatics Journal

Health Informatics Journal
2018, Vol. 24(2) 115–124

DOI: 10.1177/1460458216658778
journals.sagepub.com/home/jhi
SAGE

Cancer patients' attitudes and experiences of online access to their electronic medical records: A qualitative study

Hanife Rexhepi
University of Skövde, Sweden

Rose-Mharie Åhlfeldt
University of Skövde, Sweden

Åsa Cajander and Isto Huvila
Uppsala University, Sweden

Abstract
Patients' access to their online medical records serves as one of the cornerstones in the efforts to increase patient engagement and improve healthcare outcomes. The aim of this article is to provide in-depth understanding of cancer patients' attitudes and experiences of online medical records, as well as an increased understanding of the complexities of developing and launching e-Health services. The study result confirms that online access can help patients prepare for doctor visits and to understand their medical issues. In contrast to the fears of many physicians, the study shows that online access to medical records did not generate substantial anxiety, concerns or increased phone calls to the hospital.

Keywords
electronic medical records, medical information, patient access, patient empowerment

Introduction

Engaging patients and their relatives to play an active role in their healthcare process is a critical element of patient-centred care, yet patients are an underused resource in the healthcare system.[1] In an effort to provide more patient-centred care, some healthcare organizations worldwide have offered patients online accesses to their electronic medical record (EMR) using a secure Internet

Corresponding author:
Hanife Rexhepi, University of Skövde, School of Informatics, P.O. Box 408, SE-541 28, Skövde, Sweden.

portal, also called e-Health portal. e-Health portals are gaining traction among healthcare institutions as mechanisms to improve the safety and quality of healthcare delivery,[2] as well as modalities to activate and involve patients to a greater degree in managing their own health.[3] e-Health portals can be defined as "applications that allow patients to access health information that is documented and managed by a healthcare institution" (p. 63).[3] There is great variation in the features and functionality of available e-Health portals. Most portals allow access to selected health information from the EMR and enable patients to perform certain administrative tasks, such as appointment scheduling and prescription refills.[3,4] Besides providing sole access to health information, e-Health portals may also offer additional services such as secure messaging between a patient and an institution.[3,4] Studies have shown that giving patients' online access to health information from, for example, their EMR can encourage them to participate in their care, to manage their health condition, increase understanding of their medical issues and improve doctor–patient communication.[1,5] This is often denoted as patient empowerment, which is a situation in which the patients' role is changing from a patronized patient to an informed and engaged patient.[6] Patient empowerment is in this article defined as "patients having the ability to understand health information and make effective use of it, as well as to gain control over and participate in a meaningful way in the disease management process in an equal partnership with healthcare professionals."[4] The definition revolves around three empowerment dimensions: *patient knowledge, patient control* and *patient participation*. There is evidence in the literature that individual's ability to access and use their online medical records serves as a cornerstone of national and international efforts to increase patient empowerment and improve health outcomes. Having access to information about personal health is seen as an important precondition for patients to make informed decisions about treatment options. It also allows patients and their families to better cope with their conditions and their implications.[7] Research has also shown that online access to medical records can increase patients' ability to prepare for healthcare visits, improve communication with healthcare practitioners and increase the accuracy of information given to healthcare providers.[8] However, there seems to be a disagreement in the literature regarding the effects of making medical records available for patients, especially when it comes to anxiety and concerns. Some studies have reported that patient-accessible medical records can generate anxiety or concerns[9] whereas others have concluded that having full access to a medical record neither decreased nor increased anxiety.[5,10,11]

Giving patients' access to their EMR is not a novel idea. Nonetheless, many still primarily rely on verbal communication between healthcare practitioners and patients. Hassol et al.[2] have in their study concluded that patients and physicians differed substantially in their preferred means of communication, with patients preferring e-mail communication for most interactions followed by in-person communication, whereas physicians preferred in-person communication followed by telephone communication (p. 512).[2] Some researchers argue that the quality of verbal communication is limited due to the lack of time when visiting a physician and due to difficulties of recalling information provided during a visit.[7,12] Therefore, it is apparent that it would be beneficial to use multiple channels for communication, including written notes, brochures and online access to information that is available in the medical record. At the same time, healthcare practitioners have had several concerns about giving patients online access to their medical record.[13]

In Sweden, the County Council of Uppsala (LUL) was the first county to introduce online access to medical records by giving all patients over 18 years of age access to their personal EMR, together with several other e-Health services in the autumn of 2012. The online medical record and other e-Health services are accessed through a national e-Health portal called "1177.se." Online access to medical records enables patients in LUL to access and read their EMR containing information on appointment bookings, medical notes, drug prescriptions, medical laboratory results, diagnoses, referrals and log lists with names of the healthcare practitioners who have accessed the record. Other interactive services such as providing a healthcare declaration, changing address information, editing information about relatives and sharing the medical record with a next to kin are also provided to the patients. In the year 2015, patients were given the possibility to annotate medical notes by attaching a comment. However, healthcare practitioners are according to terms not required to keep track and/ or read any of these comments. Sweden is not the only country that is providing patients with online access to their EMR. Since 2003, patients in Denmark have had access to their EMR through a national e-Health portal called www.sundhed.dk. Even Estonia provides their citizens with access to their full personal health records. Physicians and patients have thus equal viewing access. Malta has recently introduced a Government portal for online access to health records. The portal is called Minu e-tervis (Engl. myHealth (www.digilugu.ee)). Patients and the doctors can choose who can access health data through this portal. However, to the best of our knowledge, there are no other countries except Denmark which offer patients access to all the information sets described above.

In Sweden, the patient accesses the online medical record using an e-ID or alternative secure login options. This is the same level of security that Swedish banks offer their customers for Internet transactions. Before patients can read their EMR, they are required to answer a question regarding what kind of information they want to access. They can choose to only read medical notes and laboratory results verified and approved by physicians, or they can choose to read all information, including notes, which have not been double-checked by their physician, and risking receiving information (including disturbing or worrying findings and diagnoses) before the physician has contacted them. This means that, for instance, cancer patients in LUL can see their test results before a scheduled appointment with a doctor or any other contact with healthcare. In spite of the asserted risks, currently 98 percent of the patients in LUL choose to read all available information, including notes that have not been checked by their physician. Although the Swedish Ministry of Health and Social Affairs has emphasized the importance of providing patients with a secure personal online access to their own medical records, an overview of relevant medicines and previous contacts with healthcare,[14] many physicians in Uppsala and in other parts of the country have expressed concerns that reviewing the medical record may worry and confuse patients, especially if seriously ill individuals such as cancer patients are given access to their records. Since the evidence and qualitative understanding of the impact regarding cancer patients' access to medical records has remained inconclusive in this research, a case study of cancer patients was conducted in the County of Uppsala in 2013–2014. The case study is a part of the Deployment of Online Medical Records and e-Health Services (DOME) research project.[15] Consequently, the aim of this article is to provide in-depth understanding of cancer patients' attitudes and experiences of online medical records, as well as an increased understanding of the complexities of developing and launching e-Health services based on a direct access to patients' EMR.

Research approach

Data were gathered using a semi-structured interview approach. So far, a large number of the studies on patients reading their EMR are based on surveys. In contrast to the earlier efforts, we could identify a lack of comprehensive qualitative understanding of how specific patients interact with the EMRs in a specific context. Unlike quantitative surveys, the qualitative research methodology approach allowed us to capture and explain what is going on in real organizations.[16] However, this approach also has some limitations. The qualitative interview approach with a focus on in-depth understanding means that the results are transferable through the readers' own interpretations to other settings. Another limitation is that the study is based on a convenience sample of patients with some apparent interest in the EMRs, because they were engaged enough to volunteer to participate in the study. In spite of these limitations, we argue that the material and chosen methods are appropriate in the context of the study as they provide relevant in-depth insights into the cancer patients' experiences and views of reading their EMRs.

The Regional Ethical Review Board in Uppsala approved the empirical study. Participants were recruited using an information leaflet that was placed in the waiting area at the Department of Oncology, Uppsala University Hospital during the summer and autumn of 2013. The sampling of participants was conducted in two groups. Patients in the first group (A) had consulted their EMR online, whereas in the second group (B) they had not used the service. Thirty patients (15 in each group) who had volunteered to participate in the study were contacted and subsequently interviewed by three researchers from the DOME project. The interviews were conducted in the patients' homes or at the premises of the Department of Oncology at the Uppsala University Hospital. The cancer patients were under treatment during the period that the interviews were conducted. They were also in different stages of their cancer. Some of the patients were diagnosed with cancer recurrence and had been ill for a longer time. Others were newly diagnosed. Several of the patients suffered from advanced cancer and were given palliative treatment. The patients were in the age between 30 and 92 years. Among the 30 patients, 9 were men. The length of the interviews was 45–60 min. The semi-structured interview approach meant that in addition to predefined questions, the researchers asked spontaneous follow-up questions. The interviews were transcribed by a professional and then analyzed by four researchers from the DOME project. A question–answer matrix was produced on the basis of the transcribed interview data for an analysis on a question-to-question basis. Significant phrases and quotes were coded in a separate document that was used in a thematic analysis. The quotations presented in this article have been directly extracted from the interview texts. Some passages have been rephrased to make them easier to understand.

Findings

In the analysis of the interview material of the two groups (patients from groups A and B) emerged thematic categories of reasons for why patients want to access their online medical records and potential benefits of providing online access. These themes are described in the sections below.

Why patients want to access their online medical records

Increased understanding of medical issues and increased sense of control. Patients from group A emphasized that they want to read their health information because they want to learn more about their health condition. Patients also emphasized that they use the EMR in order to know whether they have understood the information from the physician correctly or not. Other patients emphasized that EMRs helped them feel more in control of their care. The feeling of control is thus achieved first when the patient is given access to test results and medical notes, regardless of whether the information is disturbing or not. The feeling of control is perceived as crucial for well-being. Hence, for some patients, access to the medical records has had an important and crucial role in the management of their disease.

Accessing test results is crucial for well-being. Patients from the interview group A reported that the ability to have direct access to clinical test results is one of the main reasons why they have chosen to read their EMR. The patients emphasized that the healthcare system is causing them considerable anxiety because they have to wait at least a couple of weeks and at the most a couple of months, before receiving the results of a laboratory test, such as a cancer diagnosis. It was also common that patients had to wait for additional days, or weeks, before receiving laboratory results from their physician. According to the patients, the delays have a negative impact on their health. Therefore, those patients who have chosen to access their EMR argue that accessing their laboratory tests before being contacted by a doctor is a promising way to reduce anxiety and other unpleasant emotions related to waiting times. According to the interviewees, having to wait for laboratory results causes much more anxiety than accessing the results through the online medical record, even if the results would be alarming. One of the patients who read her EMR explains:

> Accessing test results, it is a tremendous difference, and it really means a lot to me. To get the information at once so you do not have to wait. It's so difficult to wait, whether it is bad or good news, it's very good to know.

All patients from group B were positive about the possibility of reading medical records online and perceive it as an important tool to increase patient engagement. In all, 13 of the 15 patients were interested in reading their medical records in the future, including information that can be worrying. The remaining patients indicated a preference to first talk with the physician. Access to information and increased patient safety appears to be two important factors that create curiosity for online medical records. The study findings also show that 14 of the 15 patients from group A, in accordance with the patients from group B, want to access all types of information, including worrying information such as cancer diagnoses. These patients argue that "to be diagnosed with cancer is worrying no matter how you get that information." Therefore, many of the patients want to decide themselves how they should receive that information, by talking to the physicians or by reading about it in the medical record. One of the patients explains:

> I think the information that you have been diagnosed with cancer is worrying no matter how you get it. […] I actually got my diagnosis by telephone, but it was my own choice that I got the information. I think that we should be free to choose how to get access to that information.

Another patient considers the following:

> If we can manage to have all these cancer diseases and to live with it, then we can handle reading about it.

Two patients from group A have received their cancer diagnosis by reading about it in their EMR, and not from talking to their physician or nurse. This was a conscious choice made by these patients. One of the patients argues that "it is easier to break down at home where you are surrounded by family, than at the doctor's office." The other patient claims that she decided to be notified about the cancer diagnosis by reading about it in the EMR as this was much easier than having to wait for information from the physician. Thus, it seems that the patients experience more anxiety when having to wait for verbal information regarding laboratory results from their physician, then accessing the results through the online medical record, even if the results would be alarming. Some patients argue:

> I'd rather sit and cry at home and fix myself in the head so I can get back on track, rather than having to sit in front of a doctor, shocked without the ability to ask questions.

> For me it was good to read about my cancer diagnosis through the online medical record. It was more difficult having to wait for information.

> I want to know even if it's bad news. It does not get any easier just because you get the same information two days later verbally from a doctor or because someone says: "it is not so dangerous and so on." No, I want clear answers.

Suspect inaccuracies. In contrast to the physicians' predictions, few patients reported being worried, confused or offended by the notes they read. Only two patients reported that they have read their EMR because they suspected incorrect entries. In all, 6 of 15 patients who had individual experiences of accessing their EMR reported that they had found inaccuracies in their medical record, and none of them, however, had filed for a correction because they did not want to be a burden for the healthcare practitioners. Moreover, all the patients who had read their EMR emphasized that when medical notes raise concerns, they usually wait to ask questions until the next patient visit instead of calling a physician. This indicates that patients are both respectful of doctors' time and resourceful in addressing questions that notes raise.

Moreover, the study findings show that 13 of 15 patients from group A have not become upset or offended after reading their medical records. When they had become upset, it was because they had found errors. However, as stated above, none of them considers the inaccuracies critical enough to require corrections. Moreover, three patients from group A argued that the access to the medical records have made the disease more evident and this has made them upset. However, they emphasize that it is not the e-Health service itself that contributed to these feelings, but the fact that they have been diagnosed with cancer. One of the patients explains: "I was upset about my cancer situation, but not for entering and reading my medical record." Moreover, one of these three patients had made an active choice to refrain from reading the medical record. Another patient, however, argues that unpleasant feelings related to the disease also occur during the patient encounter when the physician gives upsetting information, such that the patient is suffering from terminal cancer. Therefore, unpleasant feelings are according to the patient not necessarily only related to situations in which patients read their online medical record.

The importance of being able to read the medical record

Better preparation for future visits. All patients from group A emphasize that access to medical records enables them to become better prepared for their doctor visits. For example, patients from group A argue that access to the medical record prepared them for the upcoming visit. The patients prepared their doctor visits by writing down questions. It seems that those patients who prepared themselves and asked questions become more actively engaged in their healthcare and were more satisfied with the patient encounter. A patient tells the following:

> When you are visiting the doctor you get quite blocked. You can't remember. Here I have the opportunity based on what I read in my medical record to write down the questions I want to ask my doctor otherwise I might not think of them during the meeting.

All patients from group B also believe that access to their online medical records can help them to prepare for a doctor's visit.

Physician–patient relationship. In all, 3 of the 15 patients from group A felt that the preparation for future visits brought a number of other benefits, including improved physician–patient communication and increased appreciation of the physician's skill. According to these patients, being prepared for a doctor's visit contributes to more efficient communication and dialog between patient and doctor, which, in turn, seems to affect the physician–patient relationship positive. Only one of the patients, however, felt that their trust for the physician has decreased after having identified inaccuracies in the record. The remaining 11 patients do not believe that the record has affected their physician–patient relationship. Another patient stresses that the preparations have enhanced the shared decision-making with her doctor.

One interesting observation is that patients from group B argue that they have a good relationship with their physician and that they receive the information they need. This is according to the patients' one of the main reasons for why they currently do not want to access their EMR.

Another interesting observation is that patients from both groups, A and B, explained that healthcare practitioners have not informed them about the opportunity to read their medical records online. Instead, many of them have received information about the e-service through newspapers.

Aiding memory. Most patients from group A and B reported that the medical record could work as an important memory aid. Some patients from group A emphasized that it is difficult to remember all the information that was conveyed during the patient encounter. Therefore, they liked having the medical record available as a reminder before and after doctor visits. Having easy access to information about personal health when and where it is needed also seems to increase the patients' feeling of safety:

> I think you get a much better mental preparation when you have the opportunity to return to your medical record instead of just relying everything on these occasional doctor visits that are so short and so confusing sometimes.

> I'm curious about my case and I think it is good to have something to go back to. When you talk to a doctor, you will not always remember everything, therefore it can be good to be able to go back to the medical records."

Improved access to information when and where it is needed. All patients from group A emphasized that having access to medical records helps them receive information in a timely manner. Particularly, the availability of information regarding test results is considered to be crucial. Also, the patients from group B emphasize that access to the medical record may improve access to information when and where it is needed.

Furthermore, 14 of 15 patients from group A argue that medical information such as test results should be made available to the patient the minute they are available for the healthcare practitioners. They also argue that delays of publishing medical notes are more acceptable than the delay in providing access to test results. However, they argue that the delay of publishing medical notes should be no longer than 3 days. One patient from group A and two from group B do not want to take part of the test results through the EMR before they have spoken to the physician.

Learning more about their medical issues. Patients from group A reported that access to their medical records helped them understand their medical issues. Especially, by reviewing the records, they have learned more about how and when the cancer started, what treatment they have received and why and what is planned for the future. They appreciated being able to keep track of the progress of their cancer and the therapies they have received. Another patient appreciates that she is able to learn more about her medical issue in "peace and quiet" when reviewing the record. The patient argues: "I want to be able to understand my illness a little better in silence."

Moreover, patients from group A found some parts of the medical records difficult to understand. This is, however, not perceived as a major problem as the patients believe that they still have a comprehensive understanding of the content. They also argue that the understanding of the content is facilitated by the fact that they usually read their EMR after a patient visit. When the content of the medical records is not understood, 13 of 15 patients use the Internet to find information and receive answers to their questions. In some cases, the patients turn to relatives and friends who have some form of healthcare professional background. Two patients mentioned that they used other sources such as dictionaries. An interesting observation is that patients did not tend to take any additional contacts with their healthcare providers to ask questions. If the patient is unable to answer their questions by using the Internet or by asking relatives and friends, they wait until the next doctor visit.

Security and privacy. The majority of the patients from group A believe that the e-Health service is reliable and find the security satisfactory. Only one patient expressed concerns that unauthorized individuals can share the information in her EMR. Since the service has been implemented with an equivalent level of security to Internet banking, most of the patients believe that this service is not less insecure than any other national e-services. There is also an underlying expectation that these services maintain a high level of security.

> I assume that the security is very high. If not, then it should not be available to patients. It must be 110% secure so that no one but me, and those who are authorized can access and read the information.

Similar to patients in group A, patients in group B expect that the level of security of the online service is high. Only one patient expressed concerns that the security might not hold the required levels and also perceives his or her own information as sensitive. In spite of the isolated concerns, the respondent from both groups, A and B, shares an attitude that their health information is not interesting for others. One patient argues: "There are no secrets in my records and it does not bother me at all if anyone else sees my records." One patient from group B is, however, concerned about unauthorized access. The patient believes that it is important to protect the privacy of individual patients and have confidence that healthcare practitioners care for protecting it.

As presented previously, when patients access their EMR, they are required to answer a question regarding what kind of information they want to access. A warning is hence displayed regarding sensitive/worrying information. In all, 13 of 15 patients had noted the warning during the login process. Some patients even experience the warning as silly as having to answer the same question at every login is perceived as frustrating.

Moreover, only one patient has chosen to share the medical records with a family member. The patient perceives the function to be important; however, from the study one can conclude that it is seldom used. Other patients argue that they do not share their medical records as they usually read their EMR together with family members. Patients from group B have similar views as patients from group B regarding this functionality. More than half of the patients are positive to the technical possibility to share their medical records with family and friends, while others believe that they want to keep the access for themselves. Those who are positive believe that it can be useful to share the records when one is old and sick. Allowing relatives to read the EMR seems to be a way for patients to involve them in their care, and consequently to increase relatives' engagement.

Discussion and future work

Accessing medical records is a controversial issue. When medical records were made accessible online for patients in Uppsala, it raised discussions and concerns of security and ethics, both in the media and among healthcare practitioners at hospitals in the region. Many worried about the workload of the healthcare practitioners and about how patients would handle the information. One of the prominent aspects of concern was that cancer patients would be accessing test results and medical notes without being able to discuss with healthcare practitioners on possible upsetting matters right away. The possibility to receive a cancer diagnosis online was seen as especially problematic. However, little was known about why cancer patients want to read their own medical records and how they manage their personal health information. The results from this study deepen our understanding of cancer patients' attitudes and experiences of online medical records. Similar to earlier research,[6,13,17] the study results confirm that online access can help patients prepare for doctor visits, which, in turn, seems to improve the communication with practitioners. It can also help patients learn more about and understand their medical issues. An interesting aspect of the results is that the study participants had not experienced the negative aspects of online access anticipated by physicians. For example, in contrast to fears of many physicians, online access to medical records did not generate substantial anxiety, concerns or increased phone calls.[5,18] In accordance with this research, it seems that patients are more respectful of doctors' time in addressing questions that the medical records raise than many doctors have assumed.[8] Moreover, although patients did find some parts of the medical record difficult to understand, they did not perceive it as problematic. The study findings support the urges that it is crucial to include the patients' perspective in the development of e-Health services, since they have experiences and opinions unknown for healthcare practitioners. Another interpretation of the results is that the healthcare sector needs to consider patients as a heterogeneous group of people, who have different needs and habits. Access to medical records is appropriate and possibly beneficial for some patients but not for all. Therefore, there is a need to consider and respect the differences between individuals and develop e-Health services that are based on the needs of the individuals. Therefore, it may be relevant to identify the characteristics of those patients who experience anxiety and to determine which patients are and are not suited for comprehensive information through e-Health services.[8] Moreover, when it comes to security and privacy issues, there are two important reasons why patients do not seem to be worried about unauthorized access to medical records: (1) patients expect that the e-Health service has a high level of security and (2) patients perceive their own information as non-sensitive. Despite the fact that their own information is not deemed sensitive, there is a recognizable expectation among patients that their patient information should only be made available to authorized healthcare practitioners.

Because of the limitations of this study, it is not possible to reach definitive conclusions about the outcomes discussed above. Further qualitative research with a larger sample size regarding patients' access (both seriously ill patients and other patient groups) to the EMR is needed. This study is, however a step toward that direction. Future research should also focus on studying in what way the information can be adapted and improved so that patients who want can become more involved gain increased knowledge and gain better control over their own healthcare, that is, to become empowered.[4] In addition, there is a need for further research on the development of related e-Health services and on the premises and methods for facilitating improved and secure communication between patients and healthcare.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) received no financial support for the research, authorship and/or publication of this article.

References

1. Ricciardi L, Mostashari F, Murphy J, et al. A national action plan to support consumer engagement via e-health. *Health Affair* 2013; 32(2): 376–384.
2. Hassol A, Walker JM, Kidder D, et al. Patient experiences and attitudes about access to a patient electronic health care record and linked web messaging. *J Am Med Inform Assoc* 2004; 11(6): 505–513.

3. Bourgeois FC, Mandl KD, Shaw D, et al. Mychildren's: integration of a personally controlled health record with a tethered patient portal for a pediatric and adolescent population. In: *AMIA annual symposium proceedings*, 14–18 November 2009, San Francisco, California.
4. Ünver Ö and Atzori W. Support users to access information and services—SUSTAINS, http://www.sustainsproject.eu/publicdocuments/2013 (2013, accessed 3 January 2015).
5. Delbanco T, Walker J, Bell SK, et al. Inviting patients to read their doctors' notes: a quasi-experimental study and a look ahead. *Ann Intern Med* 2012; 157(7): 461–470.
6. Ammenwerth E, Schnell-Inderst P and Hoerbst A. Patient empowerment by electronic health records: first results of a systematic review on the benefit of patient portals. *Stud Health Technol Inform* 2011; 165: 63–67.
7. Gravis G, Protiére C, Eisinger F, et al. Full access to medical records does not modify anxiety in cancer patients. *Cancer* 2010; 117(20): 4796–4804.
8. Wals JS, Businger A, Gandhi TK, et al. Implementing practice-linked pre-visit electronic journals in primary care: patient and physician use and satisfaction. *J Am Med Inform Assoc* 2009; 17: 502–506.
9. Ross SE and Lin CT. The effects of promoting patient access to medical records: a review. *J Am Med Inform Assoc* 2003; 10(2): 129–138.
10. Fisher B and Britten N. Patient access to records: expectations of hospital doctors and experiences of cancer patients. *Br J Gen Pract* 1993; 43(367): 52–56.
11. Walker J, Leveille SG, Ngo L, et al. Inviting patients to read their doctors' notes: patients and doctors look ahead: patient and physician surveys. *Ann Intern Med* 2011; 155(12): 811–819.
12. Krasniqi H, Åhlfeldt R-M and Persson A. Patients' experiences of communicating with healthcare—an information exchange perspective. In: *Proceedings of the 15th international symposium on health information management research (ISHIMR)*, Zurich, 8–9 September 2011, pp. 241–251. J R Collis Publications.
13. Erlingsdottir G and Lindholm C. Because we can—eHealth encounters New Public Management in Swedish Healthcare. In: *Proceedings of the Nordic association of business administration conference (NFF)*, Reykjavik, 21–23 August 2013, pp. 1–13. Nordiska Företagsekonomiska Föreningens.
14. Ministry of Health and Social Affairs. National Strategy for eHealth Sweden, http://www.government.se/ (2006, accessed 10 January 2015).
15. DOME. Deployment of online medical records and other eHealth services, http://www.it.uu.se/research/hci/dome/index.php?lang=1 (2015, accessed 10 January 2015).
16. Jabar MA. An investigation into methods and concepts of qualitative research in information system research. *Comput Inf Sci* 2009; 2(4): 47–54.
17. Gladwin J. Increasing access to patient records. *Nurs Times* 2013; 109(35): 22–24.
18. Davies P. Should patients be able to control their own records? *BMJ* 2012; 345: e4905.

Health Informatics Journal

Article

Health Informatics Journal
2018, Vol. 24(2) 125–135

DOI: 10.1177/1460458216660754
journals.sagepub.com/home/jhi

Healthy and wellbeing activities' promotion using a Big Data approach

Diego Gachet Páez, Manuel de Buenaga Rodríguez, Enrique Puertas Sánz, María Teresa Villalba and Rafael Muñoz Gil
Universidad Europea de Madrid, Spain

Abstract
The aging population and economic crisis specially in developed countries have as a consequence the reduction in funds dedicated to health care; it is then desirable to optimize the costs of public and private healthcare systems, reducing the affluence of chronic and dependent people to care centers; promoting healthy lifestyle and activities can allow people to avoid chronic diseases as for example hypertension. In this article, we describe a system for promoting an active and healthy lifestyle for people and to recommend with guidelines and valuable information about their habits. The proposed system is being developed around the Big Data paradigm using bio-signal sensors and machine-learning algorithms for recommendations.

Keywords
Big Data, cloud computing, elderly, Internet of things, sensors

Introduction

The aging population and the increase in people with chronic diseases is a common scenario in developed countries. According to the World Health Organization, chronic diseases are the leading cause of death worldwide, as they cause more deaths than all other causes together and affect more people of low and middle income. While these diseases have reached epidemic proportions, they could be reduced significantly by combating the risk factors and applying early detection; the indoor and outdoor monitoring joined with prevention measures and a more healthy lifestyle can help so that millions of lives would be saved and to avoid untold suffering. For both chronic and pre-chronic people, several dangerous clinical situations could be avoided or better monitored and managed with the participation of the patient, their caregivers and medical personnel. This requires research and information gathering about socio-economic and environmental factors, dietary

Corresponding author:
Diego Gachet Páez, Universidad Europea de Madrid, Villaviciosa de Odón, 28670 Madrid, Spain.
Email: diego.gachet@universidadeuropea.es; gachet@uem.es

impact and life habits using sensors and devices, including software applications for monitoring personal activities and health signs. However, the use of recommender systems to promote a healthy lifestyle and wellbeing improves interaction with healthcare professional for better disease management.

The rapidly growing popularity of healthcare and activity monitoring applications for smart mobile devices like smart phones and tablets provides new ways to collect information about people's health status, both manually and automatically. Also, there are appearing new COTS (*Commercial Off-The-Shelf*) wearable medical sensors that can easily connect with smart phones or tablets via Bluetooth and transfer the sensing measures directly to a public or private cloud infrastructure. This has provided a more efficient and convenient way to collect personal health information, like blood pressure (BP), oxygen saturation, blood glucose level, pulse, electrocardiogram (ECG) and so on, that can be analyzed for generating alarms or furthermore; it would also be possible to track the patient's behaviors on a real-time basis and over long periods, providing a potential alert for signs of physical and/or cognitive deterioration.[1]

It is beyond doubt that an active lifestyle can improve health conditions;[2] a person with a poor physical health condition has a high risk of developing a chronic disease, such as diabetes or cardio-vascular disease (CVD); nowadays, advances in information and communication technologies (ICT), as for example, mobile communications, wearable computing, cloud and Big Data infrastructures make it more easy to develop integrated and cost-effective solutions for providing people feedback about their general health status and avoiding reaching a chronic disease, for example, 21 percent of the population in the United States uses some kind of sensor to track aspects related to their health, as for example, the level of glucose in blood, weight, physical activity, BP, or calorie consumption. An appropriate online track of such data could prevent consultation assistance or prevent the start of diseases resulting from unhealthy habits. Although there are a large number of wearable devices available in the market, the ways data are managed are not standard. Each manufacturer tends to have its own platform in the cloud for storing and analyzing the information that are usually also closed platforms. All this is a problem when it comes to analyzing the information, as we have to deal not only with the large amount of data but also with the heterogeneity and difficulty of access. In a Big Data scenario, the vast amount of information and the speed the data are generated, stored and analyzed have to be taken into account when designing the way the problem is going to be approached.

Related work

Empowering citizens to manage their own health and diseases can contribute to having a more effective utilization of health services, reducing costs and offering improvements in the quality of life in general. Continuous health monitoring may offer benefits to people with diseases (e.g. with chronic diseases like CVD or Chronic Obstructive Pulmonary Disease (COPD) and in need of some monitoring in their treatment), and also for healthy people, which can maintain their good state of health, preventing diseases thanks to the modification of their behavior and adoption of new healthy habits.

Some recent solutions focus on the use of sensors and bio-signals, such as devices for the measurement of BP, heart rate, pulsioximeters, physical movements and distance walked that are calculated using accelerometers or other mechanisms. These new systems are aimed mainly at users with chronic diseases, which need a deeper monitoring in some cases. Other systems are more oriented to healthy people who use them to improve their health status, or as prevention by incorporating healthy habits, favored with this kind of technology. There are also systems that offer a broad spectrum of functionalities and services and may be suitable for both types of users. The research effort in this direction has increased over the past years, and in the current European R&D Programme Horizon 2020, for example, the thematic priorities in “Health Demographic Change and Wellbeing” play an important role in the framework of the program of Social Challenges.[3]

The latest projects with similar orientation, and previously in the framework of FP7, have a significant focus on the development of these types of solutions based on ICT to promote self-management of health and disease. In some cases, the projects focus on remote monitoring and self-management of diseases such as CVD and COPD; CVD systems have a major focus on providing patient services that allow the patient to manage his or her health and be monitored with guarantees at home, without having to go as far as possible to the health center or hospital.[4] This approach is pursued in long monitoring treatments for life, and also in situations of short monitoring periods after discharge from the hospital, after a serious episode. The systems are oriented for persons who have had more acute episodes of major CVD diseases (e.g. congestive heart failure, myocardial infarction and chronic heart disease), and also to other CVD without traumatic episodes, but which otherwise tend to affect an even higher percentage of the population (as prehypertension, reaching a percentage of the population above 30%), and with a significant focus on earlier detection and prevention of the progression of the disease to more severe states. This preventive approach especially combines improvements in quality of life and cost-effective approaches to health care.

On the other hand, the COPD systems also include facilities for monitoring patients at home, trying to avoid the need to travel to hospital for maintenance.[5] Systems for monitoring include both questions to answer by the patient about symptoms daily and monitoring of health data that can be obtained by specific sensors or measuring devices (e.g. pulse rate, blood oxygen saturation and frequency respiratory) that the patient may have at home, also seeking to minimize the cost to the patient in terms of disruption in the activities of daily living. As COPD has as main cause smoking or being exposed to smoke (between 20% and 25% of smokers develop the disease), in the systems that address a preventive approach, to include facilities to quit smoking can play an important role. Thus, to support smoking quitting is a specific challenge from the point of view of prevention, and also for maintenance of more advanced COPD patients whose treatment also includes this issue as relevant.

Other projects put the focus on promoting healthy behaviors in individuals, without health problems, with a more general preventive character.[6] Recent research projects with this kind of orientation used to also have a special focus on the aging population. As representative, City4Age[7] defines a model to provide sustainability and extensibility to the offered services and tools by addressing the unmet needs of the elderly population in terms of detecting risks related to health-type problems, but also stimulating and providing incentives to remain active, involved and engaged, contributing to the design and operation of the ultimate Age-friendly City. My-AHA project[8] proposes early detection and intervention as crucial in sustaining active and healthy aging (AHA), preventing possible frequent problems such as cognitive decline, physical frailty, depression and anxiety, social isolation and poor sleep quality.

Companies like Corventis, MicroStrain, Lord or AT4Wireless are focused on the development of services and/or biomedical devices related to telemedicine and monitoring of chronic patients based on the use of data retrieved from biosensors. Other products and services are more oriented toward the monitoring of healthy activities and prevention (e.g. hours of sleep and daily walk distances to prevent prehypertension) and are offered by companies like FitBit, Jawbone or Garmin. Finally, other manufacturers like Tactio or iHealth offer products that are suitable for both types of users. Smartphone with sensors are commonly being combined with an app to process data, interpret the signals and show statistics to users. These apps carry out simple processing, so the functionality is sometimes extended by transferring the data to the cloud. These data can be processed by servers with complex algorithms. Another option is to use specific wearable tracking devices. Generally, these devices do not have a user interface for processing and showing data. So data must

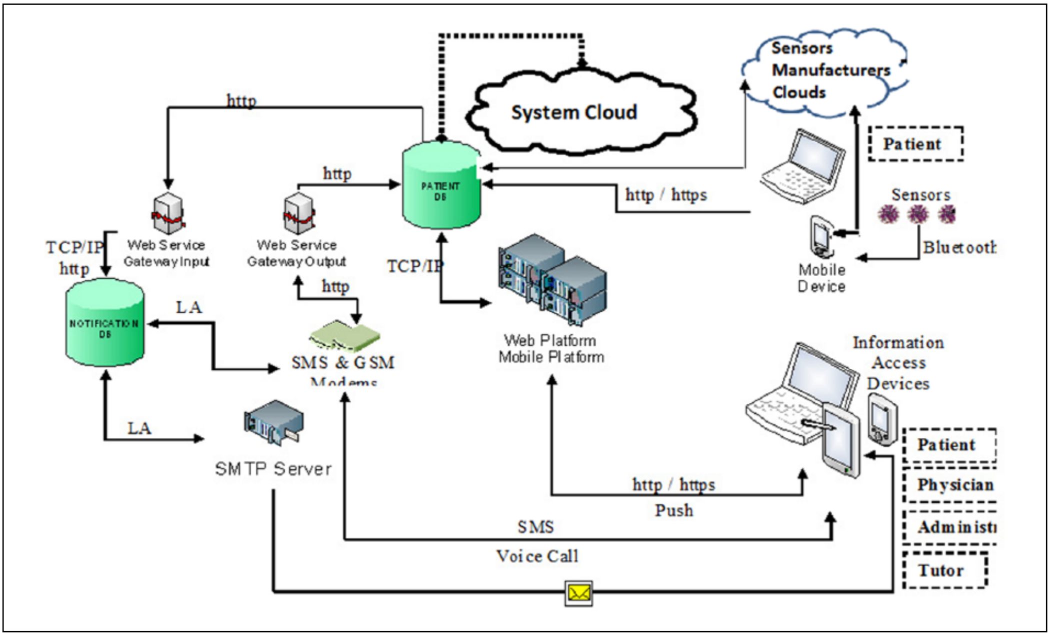

Figure 1. Proposed architecture for patient's remote monitoring.

be sent to another system to be stored, usually a mobile device, to be analyzed and to show the results to users.

Proposed architecture

The proposed architecture for collecting data in order to promote wellbeing and physical activity is based on the need for a scalable data storage and high-performance computing infrastructure for efficiently storing, processing and sharing health and activity sensor data. With this situation in mind, we propose a simple, coherent, activity monitoring solution that takes into account several factors like using non-invasive sensors, allowing the processing of high volumes of data coming from them including information from other sources as for example clinical texts; search and retrieval of medical related information from forums and designing appropriate visualization interfaces for each user type (patients, healthcare professionals, caregivers, relatives, etc.) among the implementation of security and ethical mechanism concerning the treatment of medical information.

According to the above features, our general architecture for activity monitoring as well as its associated services is presented in Figure 1. The components shown are being developed under the project ipHealth;[9] the architecture allows monitoring of both chronic or non-chronic patients and healthy people who need to be monitored by different circumstances in both home and external environments, and moreover, it allows interaction with their family, the emergency systems and the hospitals through the application of cloud computing, Big Data and Internet of things approaches. From the technological point of view, the architecture consists of the following main elements:

1. A smart mobile phone being used by user and which in turn accepts the data from wearable vital signs or activity sensors and sends this information to Internet via the mobile network 3G/HDSPA or Wi-Fi connection using sensor's proprietary applications running in the

mobile device. Sensors establish communication with the mobile device through a Bluetooth connection.

2. A cloud-based (public as Amazon Web Service or private) infrastructure for data storage and analytic module for activation of alarms to be sent to the patient and/or patient's caregivers, nursing or medical personnel. In fact, this is the core of the system in order to produce alarms or early diagnose and produce new e-health services based on the analysis of historical data using Big Data approach.[10]
3. Interoperability and messaging platform for delivery of information to all involved actors in the system, using the latest technological advances in communication (SMS, mail, voice automated systems and PUSH technology).
4. A website platform that allows both medical personnel and family caregivers to consult the associated patient information from desktop computer as well as from mobile devices.

For our system, health and activity data are taken mainly from sensor's manufacturer's clouds using open application programming interfaces (APIs) that allows developers to establish a connection between applications and health data generated by users with their products; it means we have to deal with those different APIs, which ends in a great heterogeneity of formats and services, a task that is difficult to manage from a more abstract and global point of view. In order to solve this problem, we also use "The Human API"[11] initiative, which aims to integrate, in a simple and convenient way for researchers, health data from many sensors and devices that are available in the market. Human API is a platform for working with health data that allows developers to retrieve information from different data sources (devices, wearables, APPs, services web and others) and allows users to share these data with other applications of their choice. Human API can also handle synchronizations with data from third-party sources; handles the administration of users to manage all their identities across all devices, integrated services and processes; and standardizes all data in through an API Rest that follows HIPAA (Health Insurance Portability and Accountability Act) regulations (for privacy and security of health data).

On the cloud side and due to the large amount of data to be processed, we have decided not to use the classical Structured Query Language (SQL) relational databases and file systems; instead, we decided to use NoSQL databases and a Hadoop ecosystem;[12] it enables us to implement machine-learning algorithms in both batch and stream processing. This ensures the possibility of implementing different types of analysis algorithms keeping a horizontal scalability. Recently, there have been reported experiences of using similar architectures on hospital environments of equivalent sizing.[13]

From the medical point of view, the architecture will allow the development of applications suitable for different scenarios, as for example, monitoring a prehypertensive patient who has just been diagnosed and for whom it is necessary to start a new set of healthy habits in conjunction with beta blockers that are usual drugs used as first option in these cases, and the treatment that has unfavorable effects on the patient's heart rate that needs to be determined. Controlling treatment adherence and drug effects using wireless BP monitor supposes a significant benefit for patients and doctors.

Wearable devices

About the wearable devices used in our architecture, despite there being an increasing number of applications using mobile devices and sensors for health care, such systems have a major disadvantage; in general, they do not offer a general architecture for data processing and analysis and also that approach does not consider major aspects like scalability and data security. Until recently,

Table 1. Considered sensors and applications.

Devices	Indoor/Outdoor application	Constant	Rank-alarm
Blood pressure sensor	Monitoring in cardio-vascular disease (Angor, AMI. Insufficiency heart, congenital heart disease, etc.) Monitor response to initial treatment or for comparing treatments	Blood pressure diastolic (TAD) and systolic (TAS). Heart rate	TAD <50 mmHg or >100 SBP <100 or >150 mmHg It depends on routine screening registration or acute complications
Wristband	Activity monitoring	Steps Calories burned Sleep periods	10,000 steps per day 8 sleep hours

AMI: acute myocardial infarction; TAD: tensión arterial diastólica; TAS: tensión arterial sistólica; SBP: systolic blood pressure.

continuous monitoring of physiological parameters was possible only in the hospital setting, but today, with developments in the field of wearable technology, the possibility of accurate, continuous, real-time monitoring of physiological signals is a reality. Table 1 summarizes the sensors that we are considering for use in our research about physical activity and cardio-vascular conditions.

At present, we are conducting tests for monitoring physical activity and cardio-vascular status using Bluetooth-enabled BP sensor and a FitBit Flex wristband. The iHealth BP7 wireless BP self-monitor[14] is an oscillometric device that can test, keep a history of measurements and share BP data and pulse rate with iOS (version 5.0 or higher) or Android mobile devices (version 3.0 or higher). Last firmware version is 1.3.5.

The method of measure is through automatic inflation. Regarding the validity and reliability of this monitor, it has received CE medical certification (Europe), as well as Food and Drug Administration (FDA) approval (USA) and ESH (European Society of Hypertension (HT)) Certification. The accuracy for pressure is ±3 mmHg and for pulse rate ±5 percent. iHealth Labs provides an open API that allows developers to interact with cloud iHealth's data. This API uses OAuth (Open Authorization) 2.0 protocol for authentication and authorization, the same as Facebook, Google or Twitter among others.

FitBit Flex[15] is a wrist monitor with a Micro-Electro-Mechanical Systems (MEMS) three-axis smart accelerometer that collects data about user's movement such as steps taken, distance walked and calories burned. FitBit Flex also has a sleep tracker that informs you about how many hours you sleep and the quality of your sleeping by counting significant movements. The sleep sensitivity can be configured to normal or sensitive. Flex also contains a vibration motor, which allows it to vibrate when alarms go off. Data collected by FitBit can be synced with an online website that provides graphs with the comparison of daily activity, including the activity of several months ago. Thus, the user can know which months are more active and which are less. Moreover, those functionalities can be optionally completed by filling out the menu each day. In this way, users can have complete information about the total calories consumed and the total calories burned to help them to balance both of them. Again a free app, FitBit, compatible with iOS, Android, Windows Phone, OS X and Windows, is available. This app enables to sync the monitor statistics with the mobile through BLE (Bluetooth Low Energy) 4.0 which supports encryption and authentication. Among the functionalities we can find the following: establish daily objectives and check the progress; if the user completes the menu each day into the app, then the calories consumed each day are known and it is compared with the burned; and finally the data are shared with your friends or family. As same as in other social networks, privacy preferences should be well configured in order to preserve data privacy.

The reliability of the device has been tested in several studies,[16] probing to be a valid device to measure energy consumed during physical activity. FitBit provides an API to integrate third-party applications getting and modifying user's data from Fitbit.com. The process begins with the registration of the new application which is given an API consumer key and secret. Applications must be authenticated using OAuth as same as iHealth. But FitBit uses OAuth 1.0 (it has 2.0 in a beta state) which has several vulnerabilities discovered.

Scenarios and use cases

In our project, we have worked on defining a set of scenarios and use cases applications of the proposed architecture. The scenarios include as primary users, the patients (or person improving healthy habits), health professionals who can keep track of their, and we have also considered medical students in their training process. The platform services presented here aim at enabling an integrated and intelligent access to related information for extracting useful knowledge in the context of the personalized medicine access. The scenarios considered for now are the prehypertension and the HT for young pregnant.

Prehypertension is the situation in which the patient has low BP, but it is advisable to introduce changes in his lifestyle preventively, such as weight reduction, smoking cessation, low diet in fat and sodium, physical activity and moderation in alcohol consumption, avoiding stressful situations and monitoring the adequacy of sleep. This type of data can be obtained from sensors in our architecture, and self-monitoring and its tracing facilitate the introduction of changes in lifestyle.

On the other hand, HT is the most common and important risk factor during pregnancy. There are experiences that have already proven useful in monitoring BP at home, providing accurate values. For example, in Denolle et al.,[17] the device used for this purpose automatically sends the data to be processed in order to alert the obstetrics when severe HT is produced. Moreover, that study showed that BP was higher in clinical visit than when tension was recorded automatically. Therefore, the use of sensors that automatically register the pressure levels in pregnant women proved to be useful as it may avoid unnecessary treatments for HT. In this scenario, the parameter being studied right now to be self-monitored is the BP. Other measures such as glucose in blood could be very interesting too because of the implications that could have on the mother and the baby in future.

Preliminary results

Data obtained from wearable sensors need to be processed; health data mining approaches are often used for tasks like prediction, detection of outliers (anomaly detection), clustering and decision making.[18] Depending on the kind of task, supervised or unsupervised learning methods are applied. Supervised learning is mainly used for prediction or decision making. The models obtained through learning algorithms are predictive "classifiers," that is, sets of rules or other types of models in which other attributes are used to predict the class of new examples with unknown classes. Unsupervised learning can be used for tasks like clustering or finding association rules among data attributes. Therefore, the use of sensors that automatically register the pressure levels in pregnant women proved to be useful since it may avoid unnecessary treatments for HT. Figure 2 shows results about the physical activity and BP of a monitored patient; from above to below and left to right, we can see current values about heart rate and BP among other personal data; the graphics show the amount of steps taken in a 30-day period, the steps in a day each 15 min, the percentage

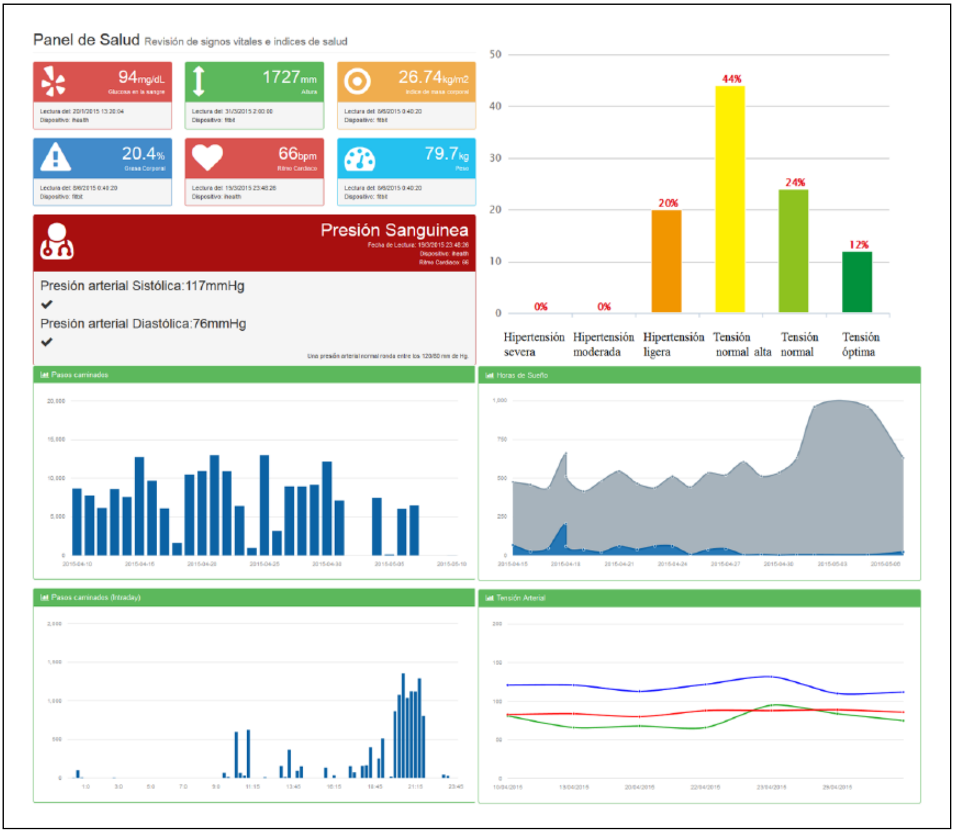

Figure 2. System user interface showing information about activity and blood pressure.

of BP range in a 1-month period, the sleep hours in a month, and the measures of BP and heart rate in 1 month.

This is valuable information for having an idea about the habits and daily patterns of a person and permits us to apply machine-learning algorithms to implement recommender systems and detect tendencies about their general health status.

Health Recommender Systems are part of Recommender Systems being applied in the health industry.[19] It has been used for diagnostic assistance by physicians and for personal health advising tools by users.[20] As the communication platform, Internet has been the main source for users to access health information and recommendations. Recommender Systems used as an enabler in health intervention bring some new functionalities,[21] for example, based on the patient's reported data, a recommender system can guide the patient about the necessary medical tests that need to be done.[22] Wiesner and Pfeifer[23] have presented a method where, by exploiting the existing semantic health network, a health-graph is constructed. The nodes of such a graph contain information that can be compared against a particular user's health-graph and essential health recommendations can be recommended. Zanker[24] has simplified the method of evidence collection by constructing rule-based preferences from historical data. According to those values, the recommender system for physical activity needs to take into account parameters and factors such as weight, age and sex. The recommendation is based on the change in BP, physical exercise and weight values of the person. Recommendations can also include the following:

1. *Dietary-hygienic measures*. Relaxation therapies depending on high BP, avoid canned food, avoid salt, reduce stress and avoid smoking.
2. *Food*. Suggestions on the best dietary changes that will improve health based on information collected in this regard.
3. *Recommendations on harmful substances*. Avoid smoking and smoky environments, alcohol.
4. *Recommendations for relaxation*. High values of heart rate during day may vary to normal levels during sleep, which may indicate stress. In this case, the recommendations include relaxing therapies like yoga. Otherwise, in case of high levels, medical assistance is recommended.
5. Sending information to healthcare or medical assistance.
6. Alerting a physician if any parameter is too high.
7. Request for consultation on the basis of combination of parameters indicating that a medical supervision is needed.

Also in our system, we consider as very important the user-generated content, like the information present in specialized forums with people with similar problems. From the point of view of the functionality in the scenarios that we have presented before, it is particularly suitable its use, allowing patients to share their concerns about certain health problems, progress associated with activities, methods to achieve goals and many other related issues with their situation of health. From the forums, users can provide indications to other relevant sources and it also allows users with professional profile to provide authorized information and monitor further comments. For the analysis of these information sources, current systems use advanced text analysis and natural language processing (NLP) techniques.

The algorithms of analysis incorporate, increasingly, the use of semantic resources, including ontologies and domain-specific databases, such as UMLS (Unified Medical Language System) or SNOMED,[25] providing a greater capacity of semantic interpretation, demanding growing computing capacity and data storage. From the point of view of automatic text analysis, our system is focused on the calculation of the similarity of consultations, giving the user direct access to the more similar that have been previously raised, finding sometimes satisfied his need or his problems resolved with previous contents. To do this, we use a similarity function[26] in a text analysis system based on Gazetteer of General Architecture for Text Engineering (GATE) and medical terms covered by the Open Biomedical Annotator (OBA) and Freebase,[27] allowing us to address synonymy and more broader lexical issues present in this domain.

Conclusion and future work

The demographic change will lead to significant and interrelated modifications in the healthcare sector and in the increasing development of technologies promoting independence for the elderly, dependents and chronic. The iPHealth project has as goal the design and development of a technological platform allowing both people with chronic diseases and healthy to increase their quality of life before more acute episodes. In this article, we have focused on the main architecture description with some details about sensors and preliminary scenarios to be considered in order to demonstrate the functionality of the developed architecture.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This work is still being developed with funds granted by the Spanish Ministry of Economy and Competitiveness under project iPHealth (TIN-2013-47153-C3-1).

References

1. Fundación Vodafone. Innovación TIC para las personas mayores. Situación, requerimientos y soluciones en la atención integral de la cronicidad y la dependencia, http://www.vodafone.es/static/fichero/pro_ucm_mgmt_015568.pdf (2011, accessed 1 February 2015).
2. Centers for Disease Control and Prevention. Physical activity and health, http://www.cdc.gov/physicalactivity/everyone/health/index.html (2014, accessed 30 May 2015).
3. European Commission. Health demographic change and wellbeing, http://ec.europa.eu/programmes/horizon2020/en/h2020-section/health-demographic-change-and-wellbeing (2014, accessed 1 April 2015).
4. Canale S. The bravehealth software architecture for the monitoring of patients affected by CVD. In: *The fifth international conference on ehealth, telemedicine, and social medicine*, Nice, France: International Society for Telemedicine and eHealth, 24 February 2013, pp. 29–34.
5. Fairbrother P, McCloughan L, Sheikh A, et al. The impact of telemonitoring COPD on continuities of care: a qualitative study. *Prim Care Respir J* 21(3): 322–328.
6. Youm S and Park SH. How the awareness of u-healthcare service and health conditions affect healthy lifestyle: an empirical analysis based on a u-healthcare service experience. *Telemed J E Health* 2015; 21(4): 286–295.
7. City4Age. Elderly friendly City services for active and healthy ageing, http://www.city4ageproject.eu/ (accessed 19 January 2016).
8. My-AHA Project. My active and healthy aging, http://www.activeageing.unito.it/ (accessed 15 January 2016).
9. ipHealth. http://www.esi.uem.es/ipHealth// (accessed 1 November 2014).
10. Gachet Páez D, Aparicio F, de Buenaga M, et al. Big data and IoT for chronic patients monitoring: Ubiquitous Computing and Ambient Intelligence. Personalization and user adapted services. In: *8th international conference, UCAMI*, Belfast, 2–5 December 2014.
11. Human API. https://humanapi.co/ (accessed 1 March 2015).
12. Prajapati V. *Big data analytics with R and Hadoop*. Birmingham: Packt Publishing Limited, 2013.
13. Sahoo SS, Jayapandian C, Garg G, et al. Heart beats in the cloud: distributed analysis of electrophysiological big data using cloud computing for epilepsy clinical research. *J Am Med Inform Assoc* 2014; 21(2): 263–271.
14. i Labs. Wireless Blood Pressure wrist monitor, http://www.ihealthlabs.com/blood-pressure-monitors/wireless-blood-pressure-wrist-monitor/ (accessed 15 January 2015).
15. FitBit. FitBit flex, http://www.fitbit.com/es/flex (accessed 15 January 2015).
16. Diaz KM, Krupka DJ, Chang MJ, et al. Fitbit®: an accurate and reliable device for wireless physical activity tracking. *Int J Cardiol* 2015; 185: 138–140.
17. Denolle T, Weber J-L, Calvez C, et al. Diagnosis of white coat hypertension in pregnant women with teletransmitted home blood pressure. *Hypertens Pregnancy* 2008; 27(3): 305–313.
18. Bellazzi R and Zupan B. Predictive data mining in clinical medicine: current issues and guidelines. *Int J Med Inform* 2008; 77: 81–97.
19. Connolly TM, Boyle EA, MacArthur E, et al. A systematic literature review of empirical evidence on computer games and serious games. *Comput Educ* 2012; 59(2): 661–686.
20. Ricci F, Rokach L, Shapira B, et al. *Introduction to recommender systems handbook*. Berlin: Springer, pp. 257–297, 2011.
21. Pattaraintakorn P, Zaverucha G and Cercone N. Web based health recommender system using rough sets, survival analysis and rule-based expert systems. In: *11th international conference of rough sets,*

fuzzy sets, data mining and granular computing, Toronto, ON, Canada, 14–16 May 2007, pp. 491–499. Berlin: Springer.

22. Ghorai K, Saha S, Bakshi A, et al. An mhealth recommender for smoking cessation using case based reasoning. In: *46th Hawaii international conference system sciences (HICSS)*, Wailea, HI, 7–10 January 2013, pp. 2695–2704. New York: IEEE.
23. Wiesner M and Pfeifer D. *Adapting recommender systems to the requirements of personal health record systems*. New York: ACM, 2010.
24. Zanker M. *A collaborative constraint-based meta-level recommender*. New York: ACM, 2008.
25. Elhadad N, Pradhan S, Gorman SL, et al. SemEval-2015 task 14: analysis of clinical text. *Semeval Notes* 2015; 298(133): 100.
26. Li Y, McLean D, Bandar Z, et al. Sentence similarity based on semantic nets and corpus statistics. *IEEE T Knowl Data En* 2006; 18(8): 1138–1150.
27. Gachet Páez D, Aparicio F, de Buenaga M, et al. Personalized health care system with virtual reality rehabilitation and appropriate information for seniors. *Sensors* 2012; 12(5): 5502–5516.

Health Informatics Journal

Article

Health Informatics Journal
2018, Vol. 24(2) 136–145

DOI: 10.1177/1460458216661861
journals.sagepub.com/home/jhi
SAGE

Electronic data capture on athletes' pre-participation health and in-competition injury and illness at major sports championships: An extended usability study in Athletics

David Karlsson and Toomas Timpka
Linköping University, Sweden

Jenny Jacobsson
Linköping University, Sweden; Swedish Athletics Association, Sweden

Juan-Manuel Alonso
Aspetar, Qatar

Jan Kowalski
Swedish Athletics Association, Sweden

Sverker Nilsson
Linköping University, Sweden; Swedish Athletics Association, Sweden

Frédéric Depiesse
French Athletics Federation (FFA), France; University Hospital of Toulouse, France; European Athletics Association (EAA), Switzerland

Pedro Branco
International Association of Athletics Federations (IAAF), Monaco; European Athletics Association (EAA), Switzerland

Pascal Edouard
French Athletics Federation (FFA), France; University Hospital of Saint-Etienne, France; University of Lyon, France

Corresponding author:
Toomas Timpka, Athletics Research Center, Department of Medical and Health Sciences, Faculty of Health Sciences, Linköping University, SE-581 83 Linköping, Sweden.
Email: toomas.timpka@liu.se

Abstract
This study set out to identify factors critical for the usability of electronic data collection in association with championships in individual sports. A qualitative analysis of electronic data collection system usability for collection of data on pre-participation health from athletes and in-competition injury and illness from team physicians was performed during the 2013 European Athletics Indoor Championships. A total of 15 athletes and team physicians participated. Athletes were found to experience few problems interacting with the electronic data collection system, but reported concerns about having to reflect on injury and illness before competitions and the medical terminology used. Team physicians encountered problems when first navigating through the module for clinical reporting, but they were not subjected to motivational problems. We conclude that athletes' motivation to self-report health data and the design of the human–computer interface for team physicians are key issues for the usability of electronic data collection systems in association with championships in individual sports.

Keywords
electronic data capture, human–computer interaction, qualitative research methods, sports epidemiology, usability

Introduction

Major international championships in individual sports, such as the International Association of Athletics Federation (IAAF) World Championships in Athletics or the International Aquatics Federation (FINA) World Aquatics Championships, attract thousands of participants and millions of spectators from all over the world. Despite the development of injury surveillance systems for use at this type of large sports events,[1] their effective implementation is still a methodological and practical challenge. It has been pointed out that the context of major sports championships features unique constraints, such as a limited window for data collection and large amounts of data to be recorded and rapidly validated.[2,3] To manage such logistical issues, electronic data collection (EDC) methods have been proposed[4–6] and introduced for use at large multi-sport events.[7] In many individual sports, overuse conditions, or tissue damage that results from repetitive demand over the course of time, are the most common health problem.[2,3,8,9] For surveillance of overuse conditions, athlete self-reporting of data on pain and other symptoms has been reported to be superior to reports from coaches or medical practitioners.[10] EDC systems have also been applied for such longitudinal self-reporting of injury data, both in Athletics and in other individual sports.[9,11]

There are ample reasons, ranging from increased data quality to efficacy, for investigating the use of EDC systems for collection of self-reported pre-participation health data from athletes and clinical data on injuries and illnesses from medical teams in association with championships in individual sports. Availability of these data sets is essential for planning, implementation, and evaluation of preventive measures. This study sets out to identify factors associated with the usability of EDC methods for gathering of data from athletes and team physicians in association with major championships in individual sports. Usability is defined in this research as the extent to which a system can be used by specified users to achieve detailed goals with effectiveness, efficiency, and satisfaction for an indicated context of use.[12] The starting point for the study focused on the users of the intended system. It assumes that the short duration of a championship differs from the longitudinal athlete surveillance setting and that the design of the EDC system has to be adjusted accordingly.[13] In system design for championship use, it also must be taken into consideration that little or no time is available for learning how to use the EDC system before starting to enter data and that data for the athlete population is culturally heterogeneous.

Methodology

This is an observational study that uses data collected from a population sample at a specific point in time. Qualitative data were collected from athletes and team medical staff using interviews and think-aloud usability evaluation methods.[14,15] Data collection was performed over a 3-day period at the European Athletics Indoor Championships during 1–3 March 2013 in Gothenburg, Sweden.

EDC for Athletics injury and illness data collection

Two demonstration versions of the online EDC systems were developed. The first was for collection of data from athletes and the second for collection of data from team medical staff. An assumption was made that there was no possibility of providing the intended users with on-site training in the EDC systems. The data collection forms in both systems were conceptually identical to the corresponding paper-based forms previously used in the injury and illness surveillance studies conducted during Athletics championships. The demonstration system for athletes (Figure 1) was developed to exhibit collection of data on pre-participation health. It was programmed for the purpose of this study and consisted of a human–computer interface for data collection without underlying database functions.

The demonstration system for use by team medical staff (Figure 2) was developed to collect data on injury and illness sustained during the championship according to standardised classifications and vocabularies.[16] This demonstration system was also programmed for the purpose of this study only, and consisted of a graphical human–computer interface without underlying database functions.

Study applications

For study of EDC on pre-participation health problems among athletes, a partial EDC system was developed using the open-source tools Node.js (https://nodejs.org/en/) and MongoDB (http://www.mongodb.org) (Figure 1). The central design goal was to provide an interface for data entry that is self-explanatory and simple to use. Evaluations of EDC systems have showed that single select questions have lower error rates compared to free text and date fields.[17] The intention was therefore to allow the athletes to complete the pre-participation form only using the mouse, that is, without having to use the keyboard. For EDC on injury and illness from team medical staff, an open-source version of the OpenClinica software package (http://www.openclinica.com) was used (Figure 2). This package is created specifically for development of EDC applications and it complies with good clinical practice (GCP) requirements. The motivation for the choice of the package was that the software both has recognised back-end security and is open-source, and thus can be progressively expanded.

Data collection

Participants were recruited by an announcement at the team medical meeting before the championships, asking athletes and team medical staff to participate. The recruitment of participants was based on the saturation principle, that is, that additional participants were invited as long as new phenomena were revealed in the data collected. Before each session, which took place at the championships venue, the participants were informed that the purpose of the systems was data collection in injury and illness surveillance studies (pre-participation and newly acquired injury and illness data) during international athletic competitions, and that no data were recorded, used and/or

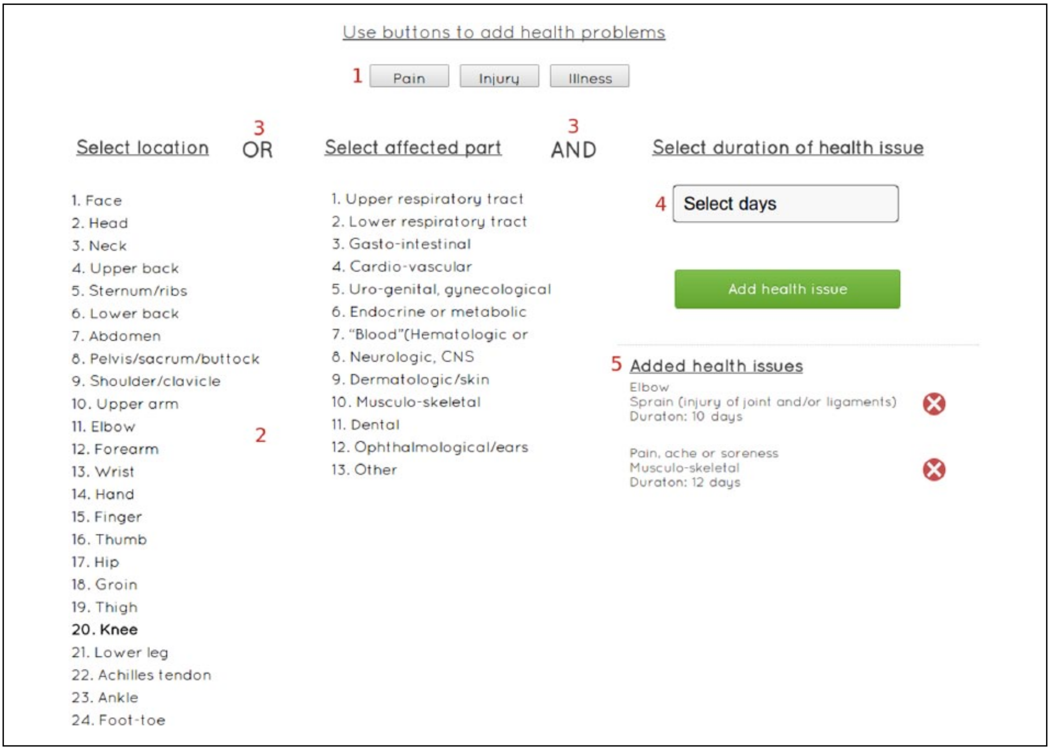

Figure 1. EDC system for athletes. Buttons are used (1) to allow selection between three general categories of health problems and (2) to display lists, different for each category. The athletes can select response items from two lists for the categories injury and illness, and from one list for pain. These selections are communicated with (3) 'OR' or 'AND'. (4) In a drop-down menu, the athletes can select the duration of health issue. The reported health issues are shown as (5) list items, which can be removed by clicking the delete icon.

diffused. Participants were then asked to complete the EDC form while thinking aloud during the process. Qualitative data were gathered from the evaluations through observations, audio recordings and field notes.[18] Each session was complemented with questions about the difficulty of the task, the usefulness of the system and opinions about reporting data online. Data saturation was found to have been achieved among athletes after eight in-depth interviews, and among team medical staff after seven interviews.

Data analysis

The evaluation data were transcribed and analysed using thematic analysis methods.[19] The focus was on identifying opinions and judgements about system use and detecting whether any features prevented task completion, caused delays, confusion or generated suggestions for improvements. Meaning units were defined as sentences containing aspects of relevance to usability and the aim of the study through their content and context. The meaning units were then coded by assigning them one or two keywords. Manifest interviewee statements (using the direct meaning without interpretation) as well as latent interpretations of statements (trying to understand to intended meaning of what was stated) about factors related to system usability were used. The codes could

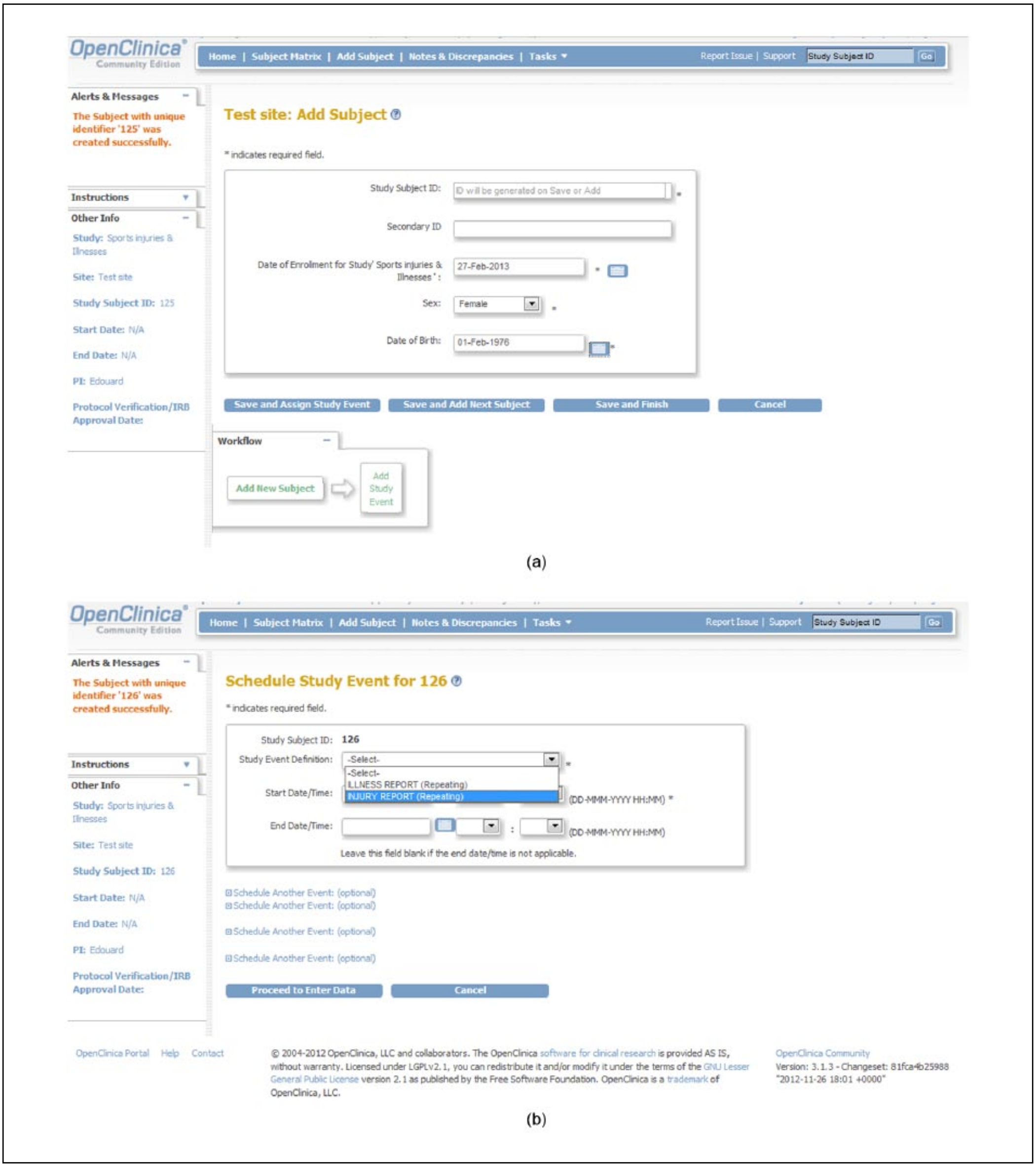

Figure 2. EDC system for team physicians. From (a) an initial display, the physician navigates to select (b) an injury or illness to report and finishes the report by navigation through a menu system. The application is intended to resemble a simple electronic medical record system.

be abstract or concrete and were used to facilitate understanding and to compare meaning units in the different system evaluation aspects that were used as categories. The categorization was first analysed by two researchers (D. K. and T. T.). Discrepancies were debated until agreement on what categories to use was reached. The categories were then organized (by D. K.) into a synthesized structure and this structure was then presented to the entire research group. In addition, the original data set was made available to the individual research group members, who were asked to make notes on suggestions for adjustments. In the final step, an online discussion was continued until consensus on the categories and category descriptions was reached.

Results

Athletes showed proficiency when completing the EDC forms for collection of data on pre-participation injury and illness but reported issues with motivation and difficulties with interpreting the survey questions and terminology. The physicians initially faced problems when navigating the multi-layered forms for injury and illness reporting. After a short learning period, they deemed completing the forms fairly easy and were generally of the opinion that EDC systems would be adequate for collecting data during Athletics championships.

EDC on athlete pre-participation health

The athletes completed the form without any human–computer interaction errors. However, the survey design posed problems for the athletes. For example, regarding the question 'How much fatigue have you experienced during the last month?' the athletes explained that they could have experienced psychological fatigue, but almost no physical fatigue or vice versa. Reporting the mean time spent on training was also found to be difficult because the athlete did not count time spent on training in weeks (but in days), which required re-calculation to get the average training time. But even those athletes who could account for hours per week also encountered problems. Some participants suggested that it would be easier to just report the time spent training for specified weeks without having to calculate the mean over a longer period. Another problem was knowing what counted as training hours. One participant said that athletes basically train 24 h a day. It was also believed that if the survey also accounted for psychological issues, then mental training should also be included in their answers. When the athletes entered data on health issues, they communicated a good understanding of previous or current injuries and illnesses, that is, the location and type of injury. The main problem with the health reporting was understanding the medical terminology used in the electronic form. Most of the athletes were of the opinion that the terminology needed to be simpler or that explanatory information was necessary. Thus, the athletes found the questionnaire and the formulation of the questions more challenging than the EDC form per se.

The main overall perception regarding the usability of the EDC system communicated by the athletes concerned their motivation to provide injury and illness data in association with championships. Even though the electronic form was found to be easy to complete, the athletes stated that they might not complete it anyway because they wanted to avoid reflecting on injuries and illnesses that might affect their performance during competition. Therefore, if they were not mandated to complete a survey before a competition, they would most likely ignore it. Some participants communicated that a positive motivational factor could be to recognise a tangible benefit for themselves. The athletes also reported that they would have preferred to fill in the form after competition.

EDC on injury and illness from team physicians

When first using the EDC system, the physicians reported concerns about navigating between the forms in the injury and illness recording system. All participants were hesitant at some point on how to continue the report. The prototype EDC system had more than 10 functions for recording injury and illness data, which meant that physicians were required to make many choices when completing tasks. For example, to initiate the report, four options were available, and to continue to the next step, it was necessary to click at 'Save and assign injury category'. All participants reported insecurity on what button to use, and one participant also clicked 'Save and finish'. The main reason communicated for this confusion was that the physicians were not familiar with the human–computer interaction terms and phrases used in the interface. It was thought that non-technical or even graphical instructions adapted to the busy context of Athletics championships would make the buttons more understandable.

For a few physicians, the navigation problems led to errors that stopped them from completing the task: that is, they clicked the wrong button and became confused about how to continue. The mechanism in these cases was clicking the wrong button while confused about how to continue to next step. However, when the physicians had received guidance and completed the task, the majority of them perceived that finishing the form had been 'fairly easy'. They reported that it was similar to the systems they usually work with in their home country. But it was also thought that too much time between reporting would make it difficult to complete due to lack of guidance in the system. Thus, contrary to the athletes, the EDC format posed more problems for the medical teams than the questionnaire itself and the terminology used.

The main overall perception communicated by the physicians was that online EDC systems would increase the efficiency of data collection by shortening the time needed for entering data. EDC was also expected to increase accessibility because paper forms sometimes were not available when an injury or illness needed to be reported. One experienced participant explained that paper forms were often unavailable at the end of championships, which could cause under-reporting of injuries and illnesses.

It was believed that an online reporting system would encourage team physicians to report injury and illness data after the championships. It was also explained that an electronic system, if available on a personal device, that is, computer, tablet or smartphone, would be beneficial. One participant suggested that a computer with Internet connection made available in the warm-up area would improve the procedure of data collection during Athletics championships. However, concerns about Internet connectivity were communicated from the participants. It was pointed out that during championships, especially if it is outdoors, Internet access is often limited and it was believed that the EDC system also needed to be available offline.

Discussion

This study set out to investigate factors critical for the usability of EDC on pre-participation health and newly acquired injury and illness in association with championships in individual sports. We observed major differences between the main user groups. Athletes reported concerns about their motivation to reflect on injury and illness before competitions, query formulations, and the terminologies used. The physicians, on the other hand, experienced problems when navigating complicated forms for recording data on injury and illness. The physicians did, however, eventually recognise EDC systems as an adequate solution for data collection during championships.

Athlete motivation was thus found to be a critical factor for success in the collection of pre-participation injury and illness data. Contemporary theories of motivation assume that people initiate and persist at behaviours to the extent that they believe the behaviours will lead to some desired outcomes or goals, for example, when they can identify a gain in personal health and well-being.[20] In this study, the athletes pointed out that they would be willing to complete the electronic form if they could identify a benefit for themselves, for example, feedback on their relative health status. Another possibility communicated by the athletes was that the coaches and medical teams demand all their athletes to complete the pre-participation forms. In other studies where athletes have been asked to provide data for registration of overuse conditions,[11] the procedure has been to first approach team coaches and ask about interest in participating in the study. Only if the coaches consent, the athletes are asked about participation. A similar approach can be tried during championships in individual sports, that is, approaching team representatives first and asking them to communicate the purposes and procedures of the study to the athletes. Nonetheless, further investigations are needed to understand how the athletes can be motivated to complete pre-participation forms, and both intrinsic and extrinsic motivation should be investigated. The second main issue associated with reporting pre-participation injury and illness data identified by athletes was the construction of the query and the terminology. These problems most likely had their origin in translations of controlled scientific notions into lay expressions, and could be at least partially addressed technically by providing online explanations.[21] Similarly, the athletes found that calculating mean training hours was difficult because the hours differed considerably between weeks, especially before competitions. They thought it would be easier to report factual training hours for defined periods of time. Such reporting could be supported technically using a drop-down menu for each period in question.

With regard to EDC on injury and illness from physicians and medical teams, we found that the data-entering task required a complex human–computer interface design, involving several levels, which made the task of entering data without previous training difficult. None of the team physicians managed to complete their tasks without receiving help and instructions. It has been shown that an EDC questionnaire that is initially too difficult to use may discourage responses and lower the quality of the data.[22] These observations suggest that the challenge of allowing medical teams to report data easily and accurately needs to be handled carefully at some level in the overall design of the EDC system. Measures can be taken either at the level of championships organisation, by scheduling instructional sessions on system use for team physicians before competitions, or at the human–computer interaction level, by ensuring that the EDC system is sufficiently self-evident in its design to allow use without separate educational interventions.

This study has both limitations and strengths that need to be considered. It addressed usability issues, and does not provide complete information for the design of an EDC system for use during championships in individual sports. However, it highlights general information system development issues that are important to consider before initiating a large-scale EDC effort. The motivational concerns among athletes and the human–computer interaction issues encountered by physicians are two examples of this. However, it is also important to acknowledge that collecting data through an EDC system is just one step in a complex workflow. Moreover, an assumption was made in the study design that the possibilities for personal on-site training on system use for athletes and team medical staff were limited. Individual training sessions on system use would probably have prevented several of the human–computer interaction issues. Although a qualitative approach based on usability tests and interviews was used, with the purpose of eliciting rich accounts of the experiences of athletes and medical teams, some important aspects and areas may have been omitted in the data. However, the similarity of the accounts collected from observed behaviours and interviews suggests that the findings can be regarded as trustworthy.

Conclusion

In this study of features critical to the usability of EDC systems for gathering data on pre-participation health and newly acquired injury and illness in association with major sports championships, we found major differences between the main user groups. Athletes' motivation to provide pre-participation data before entering the competition and the team physicians' ability to rapidly learn the human–computer interface design were key issues. We contend that a successful implementation of EDC during major sports championships is likely to require parallel adjustment of administrative and organisational processes and reallocation of resources. The results of this study can be used as a basis for implementation and evaluation of prototype systems at future championships.

Acknowledgements

The study design did not include registration of personal health data or sensitive information. Therefore, according to Swedish legislation, it did not require approval by a research ethics committee.

Declaration of conflicting interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This study was supported by research grants (FO2014-0048, P2014-0167) from the Swedish Center for Sports Research (CIF). The sponsor had no role in study design; in the collection, analysis and interpretation of data; in the writing of the report; or in the decision to submit the article for publication.

References

1. Junge A, Engebretsen L, Alonso JM, et al. Injury surveillance in multi-sport events: the International Olympic Committee approach. *Br J Sports Med* 2008; 42: 413–421.
2. Alonso JM, Junge A, Renström P, et al. Sports injuries surveillance during the 2007 IAAF World Athletics Championships. *Clin J Sport Med* 2009; 19(1): 26–32.
3. Alonso JM, Edouard P, Fischetto G, et al. Determination of future prevention strategies in elite track and field: analysis of Daegu 2011 IAAF Championships injuries and illnesses surveillance. *Br J Sports Med* 2012; 46: 505–514.
4. Bjorneboe J, Florenes TW, Bahr R, et al. Injury surveillance in male professional football; is medical staff reporting complete and accurate? *Scand J Med Sci Sports* 2001; 21: 713–720.
5. Engebretsen L, Soligard T, Steffen K, et al. Sports injuries and illnesses during the London Summer Olympic Games 2012. *Br J Sports Med* 2013; 47: 407–414.
6. Edouard P, Hertert P, Branco P, et al. Injuries and illnesses during the 2011 Paris European Athletics Indoor Championships. *Scand J Med Sci Sports* 2013; 23(4): e213–e218.
7. Derman W, Jordaan E, Blauwet CA, et al. Illness and injury in athletes during the competition period at the London 2012 Paralympic Games: development and implementation of a web-based surveillance system (WEB-IISS) for team medical staff. *Br J Sports Med* 2013; 47(7): 420–425.
8. Jacobsson J, Timpka T, Kowalski J, et al. Prevalence of musculoskeletal injuries in Swedish elite track and field athletes. *Am J Sports Med* 2012; 40(1): 163–169.
9. Clarsen B, Krosshaug T and Bahr R. Overuse injuries in professional road cyclists. *Am J Sports Med* 2010; 38(12): 2494–2501.
10. Schiff MA, Mack CD, Polissar NL, et al. Soccer injuries in female youth players: comparison of injury surveillance by certified athletic trainers and internet. *J Athl Train* 2010; 45(3): 238–242.
11. Ekegren CL, Gabbe BJ and Finch CF. Sports injury surveillance systems: a review of methods and data quality. *Sports Med* 2016; 46(1): 49–65.
12. ISO/TR 16982:2002. Ergonomics of human-system interaction – usability methods supporting human-centred design.
13. Shneiderman B and Plaisant C. *Designing the user interface: strategies for effective human-computer interaction*. 5th ed. Reading, MA: Addison-Wesley, 2010.
14. Ericsson K and Simon HA. *Protocol analysis: verbal reports as data*. Rev. ed. Cambridge, MA: MIT Press, 1993.

15. Kuusela H and Paul P. A comparison of concurrent and retrospective verbal protocol analysis. *Am J Psychol* 2000; 113(3): 387–404.
16. Timpka T, Alonso JM, Jacobsson J, et al. Injury and illness definitions and data collection procedures for use in epidemiological studies in Athletics (track and field): consensus statement. *Br J Sports Med* 2014; 48(7): 483–490.
17. Walther B, Hossin S, Townend J, et al. Comparison of electronic data capture (EDC) with the standard data capture method for clinical trial data. *PLoS One* 2011; 6(9):e25348.
18. Nielsen J. *Usability engineering*. San Francisco, CA: Morgan Kaufmann, 1994.
19. Braun V and Clarke V. Using thematic analysis in psychology. *Qual Res Psychol* 2006; 3(2): 77–101.
20. Deci EL and Ryan RM. The 'what' and 'why' of goal pursuits: human needs and the self-determination of behavior. *Psychol Inq* 2000; 11: 227–268.
21. Breakey VR, Warias AV, Ignas DM, et al. The value of usability testing for Internet-based adolescent self-management interventions: 'Managing Hemophilia Online'. *BMC Med Inform Decis Mak* 2013; 13: 113.
22. Caro JJ Sr, Caro I, Caro J, et al. Does electronic implementation of questionnaires used in asthma alter responses compared to paper implementation? *Qual Life Res* 2001; 10(8): 683–691.

Health Informatics Journal

Article

Emotional states recognition, implementing a low computational complexity strategy

Health Informatics Journal
2018, Vol. 24(2) 146–170

DOI: 10.1177/1460458216661862
journals.sagepub.com/home/jhi
SAGE

Adrian Rodriguez Aguiñaga
Instituto Tecnológico de Tijuana, Mexico

Miguel Angel Lopez Ramirez
Instituto Tecnológico de Tijuana, Mexico

Abstract

This article describes a methodology to recognize emotional states through an electroencephalography signals analysis, developed with the premise of reducing the computational burden that is associated with it, implementing a strategy that reduces the amount of data that must be processed by establishing a relationship between electrodes and Brodmann regions, so as to discard electrodes that do not provide relevant information to the identification process. Also some design suggestions to carry out a pattern recognition process by low computational complexity neural networks and support vector machines are presented, which obtain up to a 90.2% mean recognition rate.

Keywords

EEG, affective computing, emotions, neural networks, support vector machines, Brodmann regions, arousal, valence

Introduction

The recognition of emotional states through a computer has been a widely studied topic in affective computing, since Rosalind W Picard proposed the basis to analyze the physiological responses of an emotion;[1] however, most of these studies have been focused on analyzing the peripheral reactions produced by emotions, such as vocal or facial expression, and some researchers suggest that these responses could be manipulated by common users and they suggest the implementation of bio-medical signals as more reliable data sources, according to the William James theory:[2] emotional states could produce disturbances in one or more of the basic human functions, such as changes in heart rate, muscle responses or temperature changes, when a person is facing a real or

Corresponding author:
Adrian Rodriguez Aguiñaga, Instituto Tecnológico de Tijuana, Blvd. Industrial s/n, Mesa de Otay, 22430 Tijuana, B.C., Mexico.
Email: adrian.rodriguez@tectijuana.edu.mx

Table 1. Related work in classification and identification techniques of emotional states and associated identification rate (IR).

Technique	IR (%)	Reference
KNN, LDA	81.25	Murugappan (2011)[4]
MPC	64.00	Yuan et al. (2007)[6]
KNN	82.27	Heraz at al. (2007)[7]
SVM	66.70	Takahashi (2004)[8]
SVM	93.50	Li and Lu (2009)[9]
SVM	77.80	Rozgic et al. (2012)[10]
NN	43.14	Attabi and Dumouchel (2012)[11]
NN	60.00	Razak et al. (2005)[12]
NN	93.30	Gunler et al. (2012)[13]

KNN: k-Nearest Neighbors algorithm; LDA: Linear discriminant analysis; MPC: Multi-way Polarity Classification; SVM: support vector machines; NN: neural networks.

an imaginary emotional stimulus. There are many bio-medical signals that could be evaluated to recognize emotional states, such as heart rate or body temperature variations; however, our interest is focused on the analysis of signals that were originated by brain activity, particularly electroencephalography (EEG) due to the fact that this technique does not require extensive technical knowledge to implement and implementation costs are relatively low.

Related work

Some of the most outstanding work in emotion recognition by an EEG signals analysis, was presented by Dr M Murugappan of Perlis University in Malaysia and by Dr Sander Koelstra at Queen Mary University of London in the UK. Dr Murugappan proposed a mathematical model to infer emotional states from EEG analysis and Dr Koelstra developed the DEAP database, which is an extensive collection of physiological signal records of emotional stimulation processes, and both works demonstrates the feasibility of the establishment of a relationship between the electrical activity in brain cortex and emotional states.[3–5] A summary of related work is presented in Table 1.

Emotions

Disambiguation

Words such as affect, feeling and emotion are commonly used synonymously; however, each of them has a very different meaning, in their etymology and in the physical and mental reactions they cause.[14–16]

- Emotions are the manifested reactions to those affective conditions that due to their intensity, move us to some kind of action with slight or intense, concomitant or subsequent, repercussions upon several organs, that can set up partial or total blocking of logical reasoning.
- Affect could be defined as a grouping of psychic phenomena manifesting under the form of emotions, feelings or passions, always followed by impressions of pleasure or pain, satisfaction or discontentment, like, dislike, joy or sorrow.[17, ii]
- Feelings are seen as affective states with a longer duration, causing less intensive experiences, with fewer repercussions upon organic functions and lesser interference on reasoning and behavior.[iii]

Emotion theories

Over centuries, philosophers, physicians and psychologists have studied affectivity phenomena by questioning their origin, their role upon psychic life, their action with regards to favoring or hindering adaptation and their neuro-physiological concomitants;[18] however as established by Scherer (Scherer (2005)) "even the simple question, of what emotions are?, hardly get the same answer".[19]

There are classical and antagonistic theories of emotions:

- Classical theories are supported by the Darwinian theory of emotions, which states that affective reactions are innate patterns designed to orient behavior and promote adaptation,[18,20,21] and by recent theories which suggest that emotions are complex phenomena initiated by a central process as result of internal or external causes, that can be observed as an organismic alteration.[19,22–29]
- Antagonistic theories are led by the Claparede theory,[14] that defines emotions as useless, non-adaptive and harmful phenomena, since according to this theory emotions are characterized by a sudden disruption of an affective balance (mostly for short episodes), with slight or intense, concomitant or subsequent, repercussions upon several organs, that can set up partial or total blocking of logical reasoning in the affected subject.

It is the definition provided by Scherer (Scherer (2005)), which best fits an engineering task, defining an emotion as: an episode of interrelated and synchronized changes in most of the organismic subsystems(note call 4 and 5), in response to the evaluation of an external or internal stimulus event.

Data sources

The lack of a standardized database is one of the main problems with carrying out a study about the physiological responses of an emotional state, since standardized databases such as the International Affective Picture System (IAPS) or the International Affective Digitized Sound system (IADS) can be only implemented to perform the stimuli processes, which implies that if two different research groups perform an experimental setup that associates the same stimulus and even the same participants, the results will vary due the environmental conditions.[30,31]

However, there are some data collections such as bu-3DFE, PhysioNet, DEAP and the Ibug project that can help with this problem.

- bu-3DFE is a collection of several physiological signal records, associated with a wide range of emotional expressions.[32]
- PhysioNet is a collection of physiological signals, time series and images, constructed to perform a behavior analysis.[33]
- The Ibug project is a collection of bio-metric data associated with affective behaviors.[34]
- DEAP (Database for Emotion Analysis using Physiological Signals) is a wide collection of biosignals generated by several specialized experimental setups under arousal and valence stimuli.[5]

Database for emotion analysis using physiological signals

is a large collection of physiological signals which are directly associated to an emotional stimulus in a multi-modal dataset for the analysis of human affective states, where the EEG and other peripheral physiological signals of 32 participants were recorded as each watched 40 one-minute

Table 2. Regions of the cortex with its associations with the cerebral cortex and its relationship with Brodmann regions.

Processes	Brodmann regions	Cortex regions
Visual	17, 18, 19, 20, 21, 37	Temporal lobe Occipital lobe
Audition	22, 41, 42	Temporal lobe
Sensory	1, 2, 3, 4, 5, 7, 22, 37, 39, 40	Parietal lobe
Motor	4, 6, 44, 9, 10, 11, 45, 46, 47	Temporal lobe Frontal lobe

• The audition processes are related to electrodes T7, T8, F7, F8, P7 and P8.
• The visual processes are related to electrodes O1 and O2.
• The sensory processes are related to electrodes CP5, CP1, CP2 and CP6.
• The motor processes are related to electrodes FC1, FCz and FC.

long excerpts of music videos. Participants also rated each video in terms of the levels of arousal, valence, like/dislike, dominance and familiarity they experienced.[5]

This is the database that we implemented for the experimental process presented in this paper, because it is to our best knowledge the most comprehensive and reliable source of data.

Data bounding methodology

This work proposes a methodology that reduces the amount of the processed data, by defining a model that establishes a correlation criteria between emotional activity and it responses in the brain cortex and excludes regions that could not provide significant performance to the recognition process.

The Brodmann regions

Dr Korbinian Brodmann sub-divided the brain cortex into regions that appeared to have micro-structural differences and associated these regions with specific cognitive functions, such as motor processing, speech, hearing or sight. Since our work is focused on analyzing an emotional process evoked by audio-visual stimuli, we are proposing that only the electrodes that are strongly related to the audition, visual, sensory and motor[vi] regions[vii] should be considered for the digital signal processing task (see Table 2).[18,35–40]

Selected electrodes

Only 15 electrodes were set as active elements, while 22 were setted as non active elements as can be observed in Figure 1 and Table 3. This provides a data reduction of 11,264 samples per second (considering a nominal sampling rate of 512 Hz), which is a significant data reduction and consequently also a computational burden reduction.[viii]

Signal conditioning

A very wide variety of phenomena could affect the performance of the analysis of physiological data, such as a wide variety of noise, or the large amount of resources required to process these types of signals.

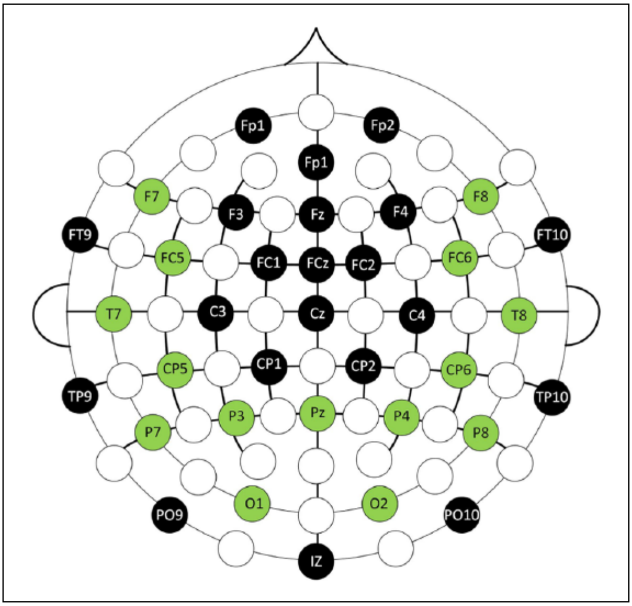

Figure 1. Graphical representation of the active and non-active electrodes (green electrodes are considered as active electrodes and black marked as non-active electrodes).

Table 3. Electrodes associated with the Brodmann regions.

Associated regions	Electrodes
Frontal Temporal 7	F7
Frontal Cortex 5	FC5
Frontal Cortex 5	C5
Frontal Cortex 6	C6
Frontal Temporal 8	F8
Parietal Cortex 5	CP5
Parietal Occipital 4	P4
Parietal Occipital 8	P8
Parietal Occipital 7	P7
Parietal Occipital 3	P3
Parietal Cortex 6	CP6
Occipital 1	O1
Occipital 2	O2
Temporal 7	T7
Temporal 8	T8

Noise filtering

The Laplacian filter described by Murugappan (equation (1)), was implemented to mitigate the problem that EEG signals are naturally contaminated with noise and artifacts (i.e. eye movement(EOG), muscular movement (EMG), vascular movements (ECG) and kinetic artifacts)[4]

$$x_{\text{new}} = [x(t) - 1/N_E]\left[\sum_{i=1}^{N_E} x_i(t)\right] \tag{1}$$

where x_{new} is the filtered signal, X_t the raw signals and N_E is the number of neighbor electrodes.

Signal bounding

A band-pass filter with cutoff frequencies of 0.5 Hz to 47 Hz, was implemented to exclude all frequencies that are not associated with the brain rhythms model: delta (0.2 to 3.5 Hz), theta (3.5 to 7.5 Hz), alpha (7.5 to 13 Hz), beta (13 to 28 Hz) and gamma ($>$ 28 Hz).[41,42]

Blind source separation

A blind source separation (BSS) algorithm was implemented to remove redundancy between active elements but preserve information of non-active elements. Since these cases are generally considered as a multi-channeling problem and the signal $y(n)$, its components $yi(n)$, could be defined as:

$$p_Y(y(n)) = \prod_{i=1}^{m} p_y(yi(n)) \quad \forall n \tag{2}$$

where $p_{(Y)}$ is the probability distribution set, $p_y(yi(n))$ is the marginal distribution and m refers to the predefined independent components, to separate each element of our gating region as shown in Figure 2, where $S_i = e_i + \sum_{m=1}^{n} r_m$, it is the sum of overlapped signals that occur when you try to read an specific electrode and $S_{bi} = e_i + \lambda_i$ represents the information of a specific an electrode without redundancy of the active electrodes.

Feature extraction

A common question about the implementation of the wavelet transformation is 'Why does it not use Fourier traditional methods?'. The answer is that there are two important differences between Fourier and wavelet analysis.

- The first is that due to the Fourier basis, functions are localized in frequency but not in time (a small frequency change in the Fourier transform could produce changes in all parts in the time domain), unlike the wavelet transform which presents resolutions in frequency (through expansion) and in time (through translations).
- The second is that many kinds of functions that can be represented by wavelets in a more compact form (i.e. functions with discontinuities and features with sharp spikes usually need fewer functions when they are analyzed on a methodology based on wavelets) and due to this, large data sets can be easily and quickly transformed by the discrete wavelet transform (the counterpart of the discrete Fourier transform) by encoding the data as wavelet coefficients, which implies a higher processing speed since the computational complexity of the fast Fourier transform is $O(n\hat{\text{A}}\cdot log2(n))$, while for the wavelet transform this is reduced to $O(n)$.[43]

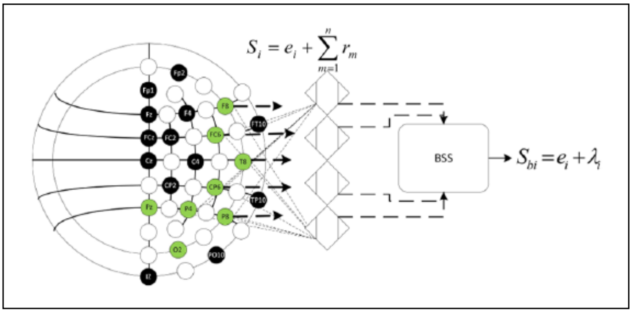

Figure 2. Process model of the blind source separation by independent component analysis implemented.

Feature selection

Each level of the discrete wavelet transform is calculated by passing the signal through a series of filters; samples are passed through a low pass filter g and simultaneously a high pass filter h, as[ix]

$$y[n] = (x * g)[n] = \sum_{k=-\infty}^{\infty} x[k]g[n-k] \tag{3}$$

Also the energy content in a specified region of the signal $f(t)$, can be decomposed as

$$f(t) = \sum_j \sum_k d_{j,k} \psi_{j,k}(t) = \sum_j f_j(t) \tag{4}$$

where j, $k \in Z$ y $\psi(t)$ is the *mother* wavelet and the coefficients $d_{j,k}$ represent the inner product (equation (5))

$$d_{j,k} = \left\langle f(t), \psi_{j,k}(t) \right\rangle = \frac{1}{\sqrt{2^j}} \int f(t) \psi(2^{-j} t - k) dt \tag{5}$$

$d_{j,k}$ represents the energy of the detail coefficients[x] from f at level j, and the classification input array can be obtained by

$$E_j = \sum_k d_{j,k}^2 \tag{6}$$

This same procedure can be implemented to extract the approximation coefficients.

The total coefficients energy can be expressed as

$$E_t = \sum_{i=1}^{e} E_j \tag{7}$$

where e, are the electrodes associated with the wavelet analysis and j is the corresponding energy percentile level calculated as

$$\varepsilon_j = \frac{E_j}{E_t} \times 100 \tag{8}$$

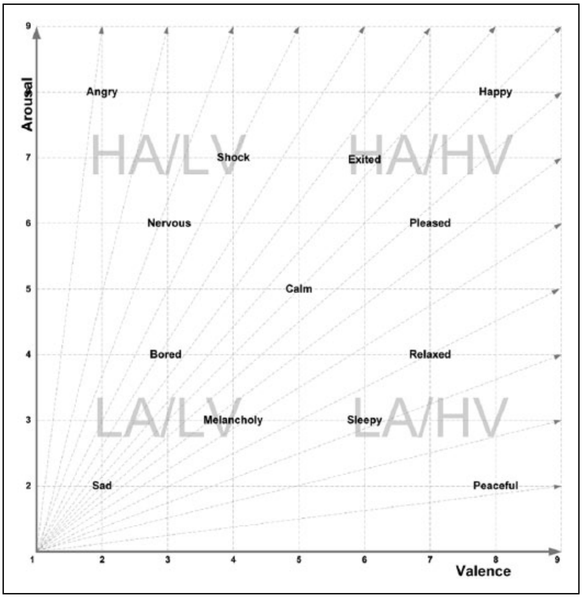

Figure 3. Emotional states distribution model, based on arousal and valence levels and the discrete tags of the Ekman model.

The translation and dilation coefficients, can be implemented directly as the features in the classification problem.

Classes

Our proposal to establish an appropriate model that establishes a clear distance between each emotional tag, is the assignment of a distance parameter between the emotional tags according to the Ekman, Russell and Scherer models. Since the Ekman model provides the basic tags to identify each class, while the Russell and Sceherer models define emotional states as arousal and valence levels.

This model can be observed in Figure 3, which encompasses all similar emotional states according to their arousal and valence levels (i.e. all emotions that have high levels of arousal and high levels of valence are correlated to states of happiness, states which have a low valence and high arousal are correlate to anger, low arousal and low valence to sadness, and low arousal and high valence to relaxation states).

- Class 1: (HA-HV) high arousal, high valence.
- Class 2: (HA-LV) high arousal, low valence.
- Class 3: (LA-HV) low arousal, high valence.
- Class 4: (LA-LV) low arousal, low valence.

Also this model defines the orientation of each of the evoked potentials according to their characteristics to define the elements of each of the classes that would be implemented as references in the identification model, as can be observed in Figures 4 and 5, where each element of the experimental process belongs to a particular class.

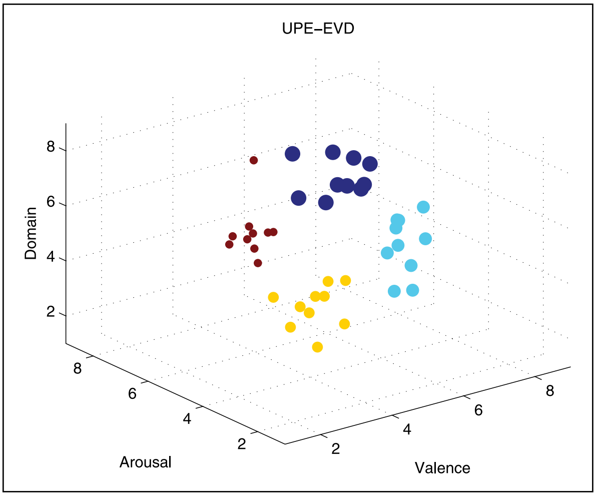

Figure 4. The location of the experimental cases, according to their emotional stimuli responses obtained by the user ratings (the input elements to the classification process are created based on this information). Arousal: it includes features that define idle or alert states (i.e. disinterested, boring, alert, excited). Valence: ranges from unpleasant (i.e. sad, stressed) and nice (i.e. happy, euphoric). Domination: ranging from a feeling of helplessness and weak (uncontrolled) to one feeling empowered (in control of everything).

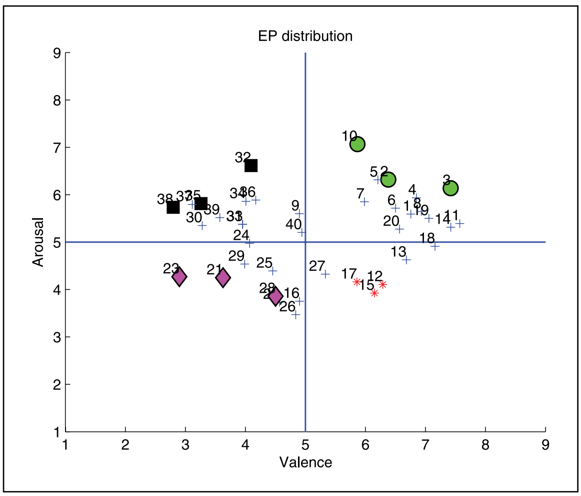

Figure 5. Distribution of evoked potential trials in an arousal and valence space. As seen there are ten experiments related to each class and a greater emphasis are given to three of them in relation to its distance to the other cases to ensure a greater separability between classes, according to their labeling.

Identification process

Support vector machines and neural networks, are the techniques that according to literature review have shown the best recognition rates (Table 1). Therefore, these two techniques were implemented in this work in order to corroborate the performance of our methodology.

Experimental setup and inputs

The experimental configuration presented in Figure 6, was designed to evaluate the identification performance of each of the combinations produced by implementing the clustering Algorithm 8.1, which ensures a consistent experimental process distribution for each of the elements of the classes by considering at least one experimental class associated with each class (i.e. each of the cross validations evaluate distinct cases of the same class).[xi]

8.1.1 Classification inputs. To ensure that each case study contains the information from more than one user at the same time and to corroborate the existence of a correlation between different study subjects, each case is also assessed by means of the following classification scheme:

- a: Contains the information of a single user.
- b: Contains information of two users.
- c: Contains information of three users.
- d: Contains information for all users.

Algorithm 1: Feature input arrangement algorithm.

Input: Features $d_{j,k}$ arrays
Output: Matrix arrangement E_k
Selecting features to provide input classes;
for $i \leftarrow 1$ **to** 3 **do**
  e= electrodes array);
  for $j \leftarrow 1$ **to** 15 **do**
    $C = [e(b); d_{j,k}]$
  'a' is the parameter that defines the number of emotional states that would be analyzed;
  for k **to** a **do**
    for $l \leftarrow 1$ **to** 30 **do**
      $E_k = [l; C]$;
  E_k *is the arrangement of features of the three case studies of each emotional state and all users*

Neural networks

Artificial neural networks (NN) are computational techniques that can be trained to find solutions, recognize patterns, classify data or forecast events by defining the way its individual computing elements are connected, which automatically adjusts their configurations to solve specific problems according to a specified learning rule. Due to the fact that EEG data are considered as chaotic signals, many researchers have proposed the implementation of NN, as one of the most appropriate tools to carry out the analysis, since NN have a remarkable ability to derive meaning from complicated or imprecise data. Besides NN are inspired by the natural behavior of a neuron, and EEG signals are signals produced by neurons.

Implementation

The implementation of a highly complex NN was our initial proposal to carry out the pattern recognition task; however, as can be observed in Figure 7, the error rates have a direct correlation to

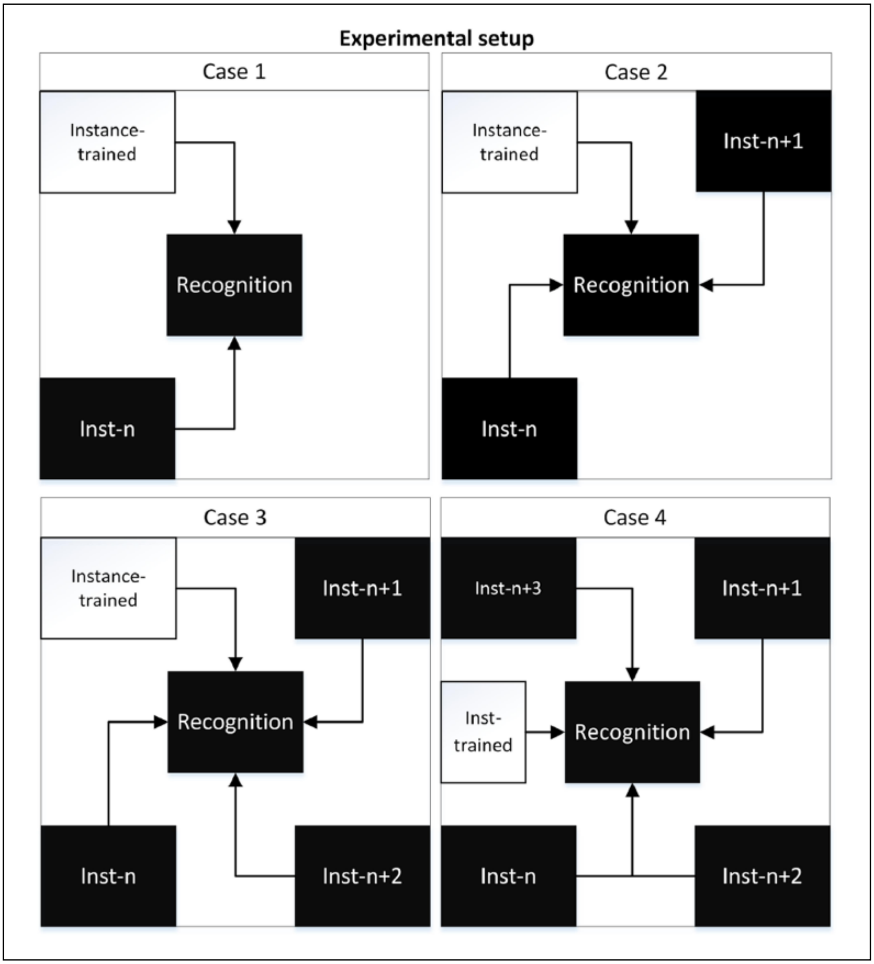

Figure 6. Class selection model to perform the identification performance task.
• Case 1: Single class training, in order to achieve identification of a particular state (n).
• Case 2: Two classes training, to carry out the identification of two emotional states (n , $n+1$).
• Case 3: Three classes training, multi-class identification for three emotional states (n , $n+1$, $n+2$).
• Case 4: All classes training for a multi-class identification (n , $n+1$, $n+2$, $n+3$).

the network architecture (i.e. if network architecture increases, so to does the error), contrary to our initial thought and that the implementation of low complexity architecture will be the best option for this particular case, as shown in Figure 8, where a 10-fold cross-validation was performed to evaluate 20 configurations of low complexity networks, in order to determine the amount of units per layer that would be necessary to obtain the most accurate and stable recognition rate.

Based on the presented considerations, the following NN was implemented to perform the presented experimental results:[xii]

- resilient back-propagation algorithm;
- 11-layer topology, 2/3/4 units in the input layer, 18 hidden units per layer and 2/3/4 units on the output;

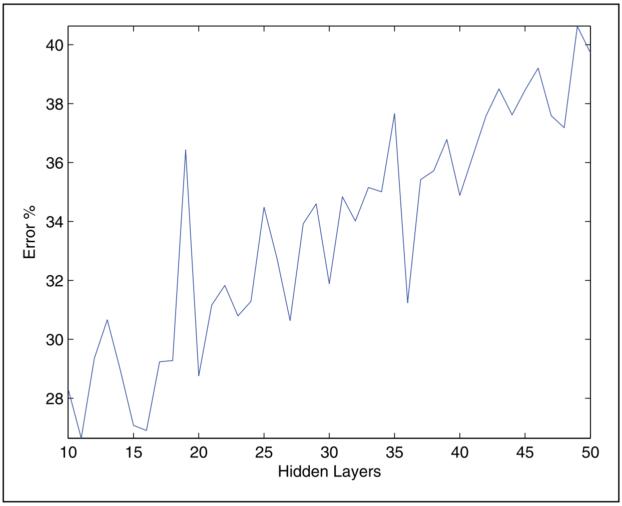

Figure 7. An identification process performed by the same NN configuration, but incrementing by one the number of layers in each iteration up to 50.

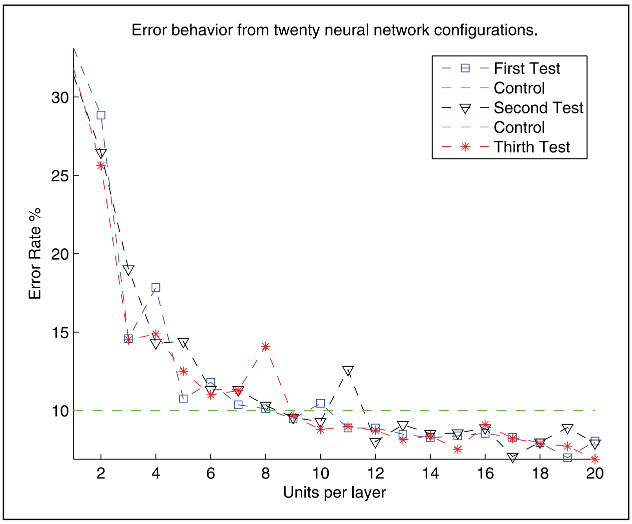

Figure 8. Performance of low complexity NN configurations. Considering only NN with less than 20 hidden layers, the validation process chose the number of units per layer this network would have by considering that 20 would be the maximum number of elements per layer.

- 200 maximum epoch;
- maximum square error (MSE), Goal = 0.002;
- 70% for training data;
- 15% to validate;
- 15% to test.

Support vector machines

Support vector machines (SVM) provide separability to non-linear regions by implementing kernel functions that avoid the *local minimum* issues by implementing quadratic optimization, so that, unlike NN, this technique is more related to an optimization algorithm rather than to a greedy search algorithm. Also when the classification problems do not present a simple separating criterion, there are several mathematical approaches that could be applied to the SVM strategy, in order to retain all the simplicity of hyperplane separation.

Implementation

Three transformation kernels were implemented to perform the SVM identification processes:

- Gaussian or radial basis $G(x,y)=\exp(-(x-y)'(x-y)/2\sigma^2))$;
- polynomial $G(x,y)=(1+x'y)^d$;
- multi-layer perceptron $G(x,y)=\tanh(p_1xy+p_2)$.

Besides the fact that the training algorithm function implements an optimization method to identify vectors s_j, weights α_j and bias b, which is used to classify the vectors x according to equation (9), the results of the training process are also considered in the optimization of the classification process. This condition is known as the condition of Karush– Kuhn– Tucker (KKT), which is analogous to the condition where the gradient must be zero or at least modified to consider its limitations as

$$c\sum_i \alpha_i k(s_i,x)+b \tag{9}$$

where k is the kernel function and for the lineal case and if c $\geqslant$ 0, x is considered as a member of the first group and otherwise as a member of the second group, also restricted by the Lagrangian conditions, as

$$L(x,\lambda)=f(x)+\sum\lambda_{g,i}g_i(x)+\sum\lambda_{h,i}h_i(x) \tag{10}$$

where $f(x)$ is the objective function, $g(x)$ is the conditioning function vector when $g(x)\leqslant 0$ and $h(x)$ is the the conditioning function vector when $h(x)=0$. λ vector, is then a concatenation of the λ_g and λ_h values of the Lagrange multiplier, with the the conditions of $\nabla_x L(x,\lambda)=0$, $\lambda_{g,i}g_i(x)=0 \quad \forall i$, $g(x)\leqslant 0$, $h(x)=0$ and $\lambda_{g,i}\geqslant 0$, to reduce the computational burden in the identification process, and the support vectors are defined by a $n_{sv}xp$ matrix, where n_{sv} is the number of support vectors (maximum size of the training sample) and p are the elements of the β vector (which is the numeric vector of linear predictor coefficients), having a length equal to the number of predictors used to train the model and α values are vectors of n_{sv} elements (which can be very large for data sets that contains many features).

Results

Separability

The first thing that may be noticed is that this methodology provides a significant reduction of the computational complexity, since, as can be observed in Figures 9 to 17, the degree of separability

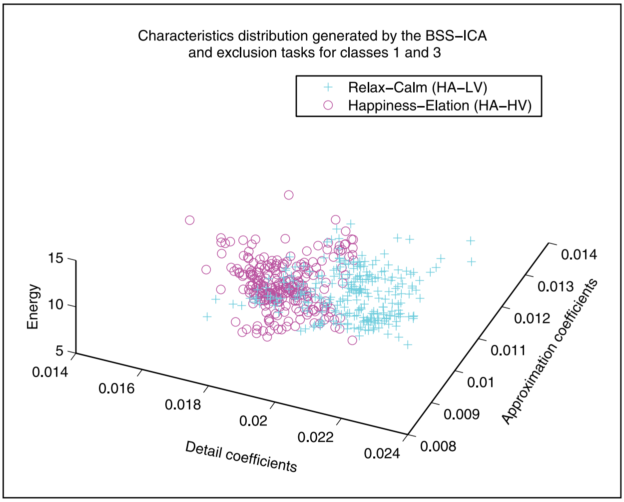

Figure 9. Distribution comparison between classes 3 and 1. A separation between the coefficients associated with these classes can be easily observed, even before the classification and recognition processes (class 3: relaxation/calm, class 1: happiness/elation).

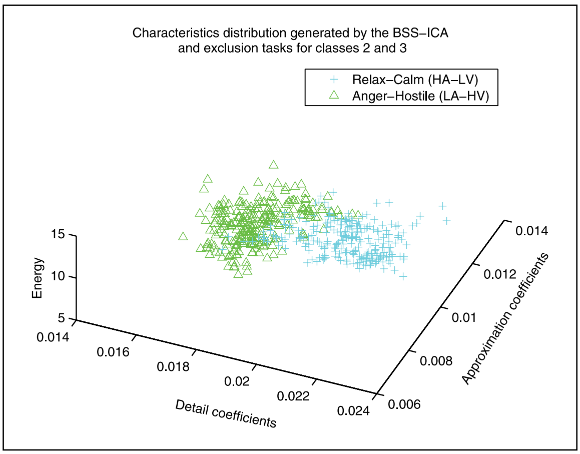

Figure 10. Distribution comparison between classes 3 and 2. A very clear separation between the coefficients associated with these classes can be easily observed, even before the classification and recognition processes (class 3: relaxation/calm, class 2: anger/hostile).

obtained by implementing it can be easily observed and some of the behavioral tendencies can be noticed without a computational identification process.

Implementation of class three as reference. Figures 9, 11 and 10 show feature distribution models that implement class 3 as reference. A very weak correlation between classes 2, 4 and 3 can be observed, while a sightly higher correlation can be observed for classes 1 and 3 (i.e. the relaxation state is more related to happiness than to anger or sadness).

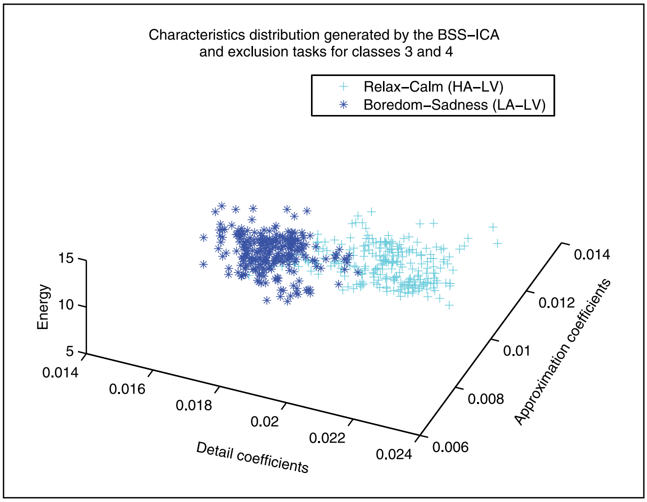

Figure 11. Distribution comparison between classes 3 and 4. A clear separation between the coefficients associated with these classes can be easily appreciated, even before the classification and recognition processes (class 3: relaxation/calm, class 4: boredom/sadness).

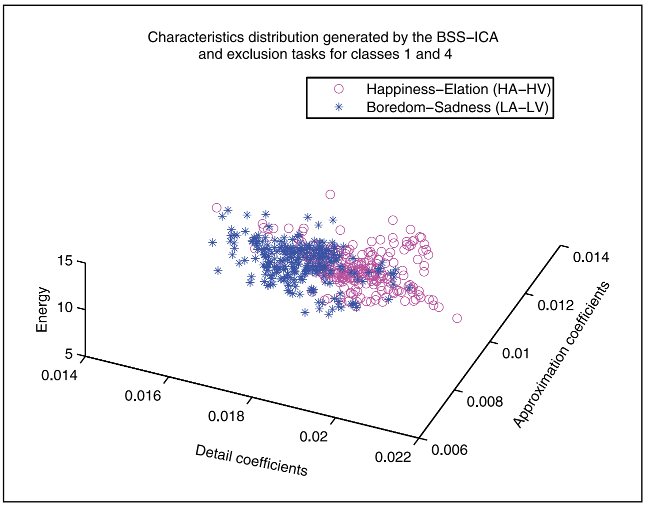

Figure 12. Distribution comparison between classes 1 and 4. For this case the separation between classes is not as clear as the three previous cases; however, that the separation process can be carried out with simple classification techniques can still be observed (class 1: happiness/elation, class 4: boredom/sadness).

Implementation of class one as reference. Figures 12 and 13 show the feature distribution models that implement class 1 as reference. It can be observed that the correlation between classes 1 and 4 could be considered weak, while the correlation between classes 1 and 2 could be considered strong (i.e. the happiness state is more related to anger than to sadness).

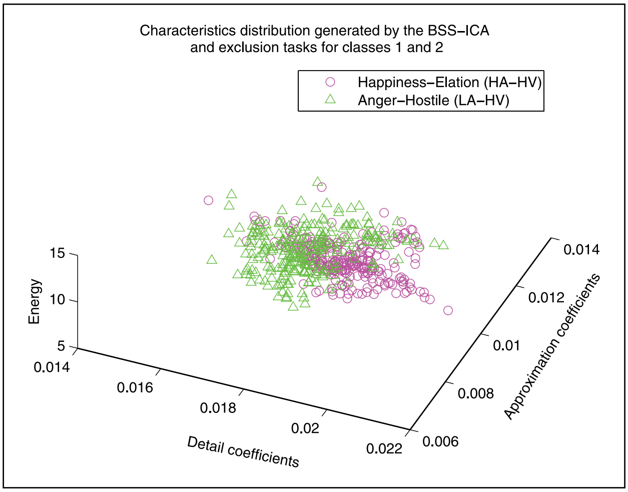

Figure 13. Distribution comparison between classes 1 and 2. For this case the separation between classes is not as clear as the three previous cases; however, that the separation process can be carried out with simple classification techniques can still be observed (class 1: happiness/elation, class 4: anger/hostile).

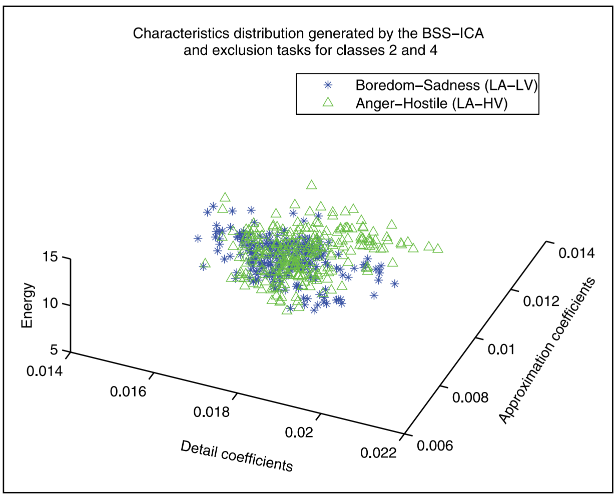

Figure 14. Distribution comparison between classes 4 and 2. For this case the relationship between the classes is considered very close, which could indicate very similar behavior in both classes (class 1: boredom/sadness, class 4: anger/hostile).

Implementation of class two as reference. Figure 14 shows a feature distribution model that provides a comparison of classes 2 and 4; the relationship between these emotional states is considered very close.

Implementation of a three-classe comparison. In Figures 15 and 16 it can be seen that even though there is clearly overlap between classes, they are mostly distinguishable to the naked eye, except for those combinations labeled with a close relationship, such as anger and sadness.

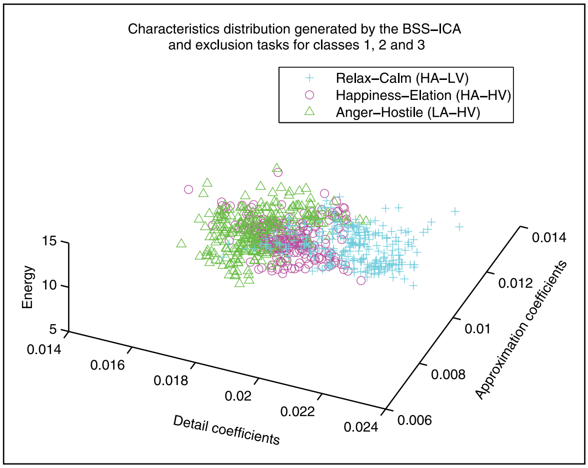

Figure 15. Distribution comparison between classes 1, 2 and 3. This representation shows the level of complexity for the three-class identification process (class 1: happiness/elation, class 2: anger/hostile, class 3: relaxation/calm).

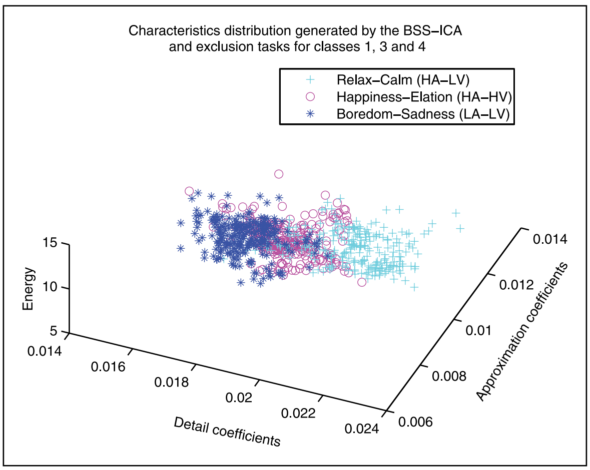

Figure 16. Distribution comparison between classes 1, 3 and 4. This representation shows the level of complexity for the three-class identification process (class 1: happiness/elation, class 3: relaxation/calm, class 4: boredom/sadness).

Implementation of a four-class comparison. Figure 17 provides an overview of the four emotion recognition process that was implemented in this work; it can be observed that these emotions are closely related and represent a considerable increase in the complexity of identification.

NN performance

The implemented NN obtains up to 98% mean recognition rate for the trivial case (by recognizing a single emotional state), while for the binary and the multi-class scheme (two, three and four emotions) the mean identification rates were up to 90.2%, 84.2% and 80.9%, respectively.

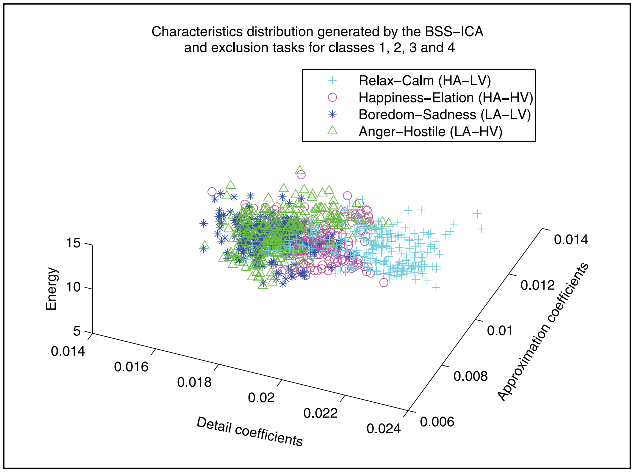

Figure 17. Distribution comparison between all classes. This representation shows the level of complexity for the four-class identification process (class 1: happiness/elation, class 2:anger/hostile, class 3: relaxation/calm, class 4: boredom/sadness).

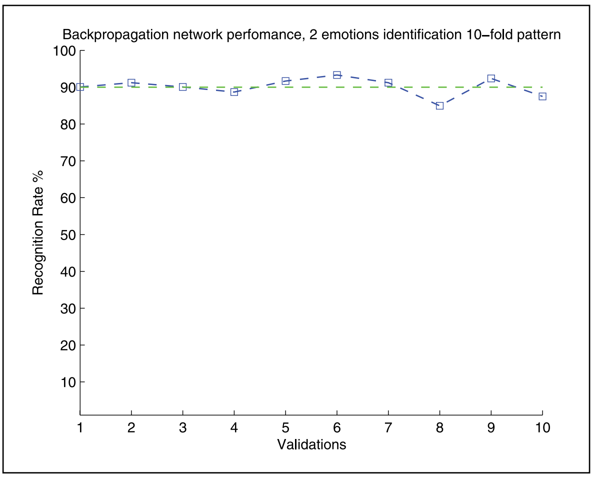

Figure 18. The 10-fold cross-validation performance of the NN for the two-class identification process.

The performance of the NN for the two classes recognition process can be observed in Figures 18 and 19, in which can be appreciated the recognition rate by class and a graph showing ten validations, showing a stable performance.

Figures 20 and 21 present the identification performance for the three-class recognition scheme, which obtains up to an 84.2% mean identification rate and the 10-fold cross-validation process to evaluate the stability of this identification scheme.

Figures 22 and 23 presents the four-class recognition scheme which obtains up to an 80.9% mean identification rate and the 10-fold cross-validation process to evaluate the stability of this identification scheme.

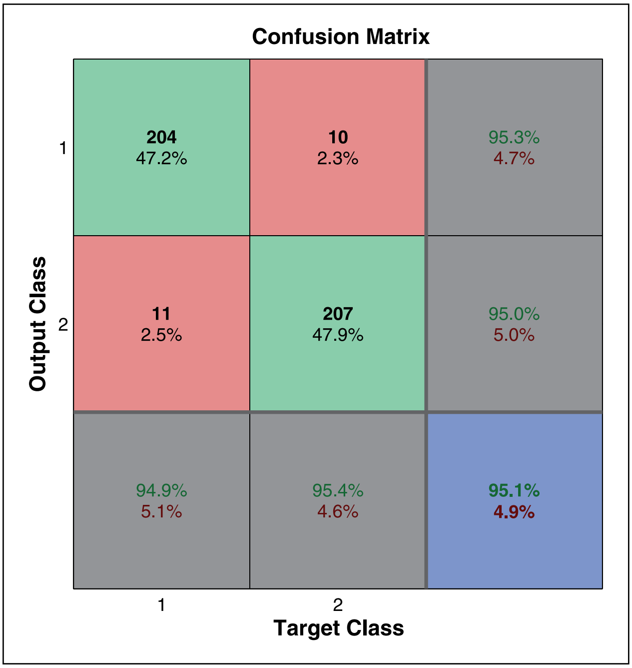

Figure 19. Performance of the NN for the two-class identification process.

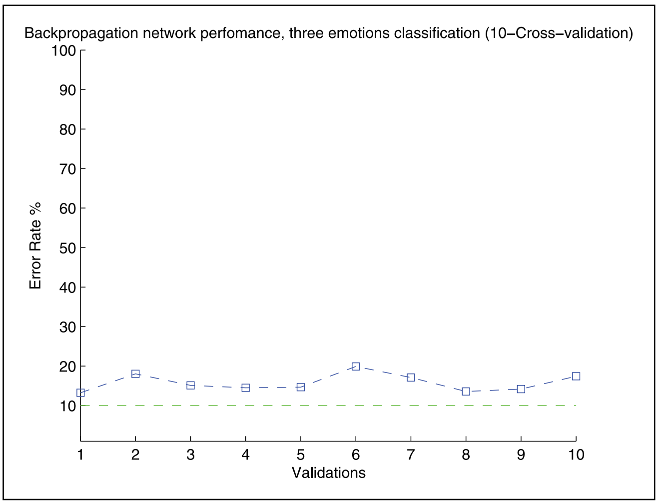

Figure 20. The 10-fold cross-validation performance of the NN for the three-class identification process.

SVM performance

The results obtained by implementing the support vector methodology and the proposed signal conditioning strategy are shown in Tables 4, 5 and 6.

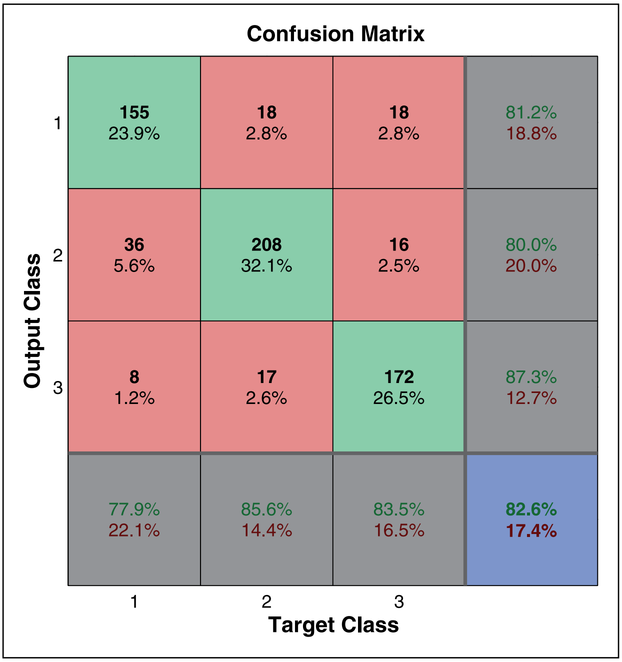

Figure 21. Performance of the NN for the three-class identification process.

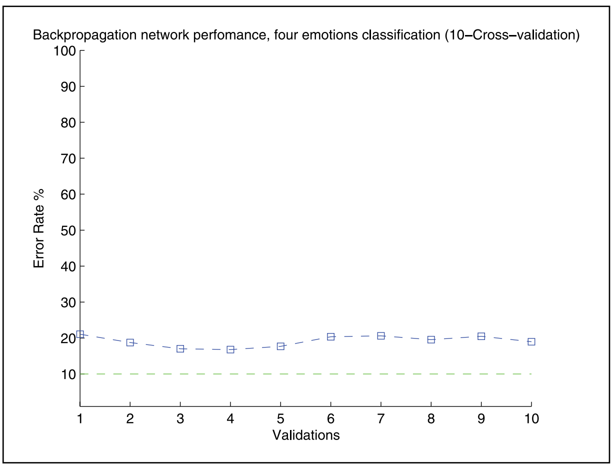

Figure 22. The 10-fold cross-validation performance of the NN for the four-class identification process.

SVM identification performance

As can be observed in Table 7, the performance obtained from implementing this identification technique and the proposed signal conditioning strategy remains in competitive range when compared with the rates reported in the literature, although it could be considered that some other authors obtain slightly higher identification rates,[9,13] despite the fact that their implementations do

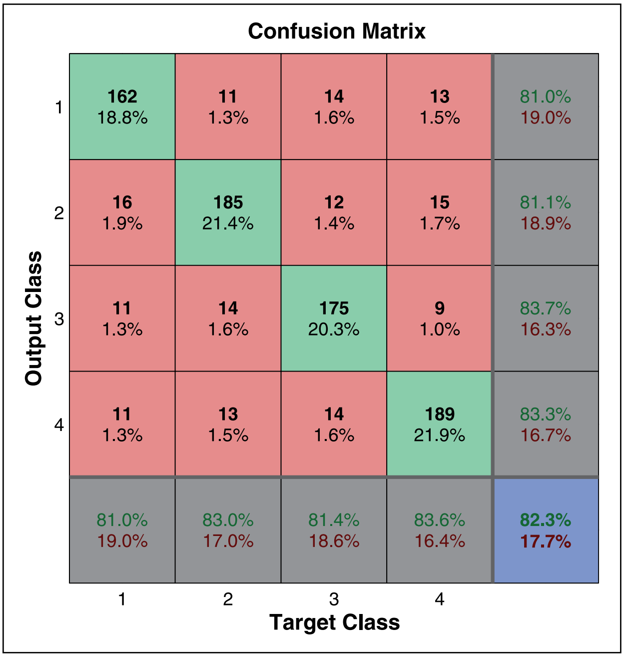

Figure 23. Performance of the NN for the four-class identification process.

Table 4. SVM average identification rates for the Gaussian experimental cases.

	Gaussian			
	Identification Rates (%)			
	Case 1	Case 2	Case 3	Case 4
a	89.59	83.61	79.20	82.61
b	88.78	83.22	81.19	80.56
c	86.61	81.22	78.53	81.23
d	84.03	84.85	83.88	83.38

not consider multi-class problems, while works that do consider multi-class schemes obtain significantly lower identification rates than those presented in this paper.[4,5]

Conclusions

We present a strategy to carry out an emotion identification process, by analyzing the electroencephalographic activity of users when they are experiencing one or more emotional stimuli by implementing a strategy that reduces the amount of electrodes that have to be analyzed, which can be translated as an a priori removal of large amounts of information from the outset, retaining competitive recognition rates. In addition, some considerations that allow for the reduction of the computational burden required for the recognition and identification process are presented.

Table 5. SVM average identification rates for the polynomial experimental cases.

	Polynomial			
	Identification Rates (%)			
	Case 1	Case 2	Case 3	Case 4
a	90.19	82.34	78.76	82.67
b	89.45	83.76	81.56	81.23
c	89.12	82.21	79.50	82.47
d	85.47	84.73	84.49	85.15

Table 6. SVM average identification rates for the multi-layer experimental cases.

	Multi-layer			
	Identification Rates (%)			
	Case 1	Case 2	Case 3	Case 4
a	89.89	83.43	78.16	81.11
b	89.85	82.98	83.12	81.64
c	87.68	82.47	77.93	80.51
d	85.48	86.64	83.13	83.18

Table 7. Performance comparison of the implemented classification techniques.

Overall Performance (%)			
Cases	Case 1	Case 2	Multi-class
SVM Gaussian	87.25	83.22	81.32
SVM Polynomial	88.55	83.26	81.97
SVM Multi-Layer	88.22	83.88	81.11
NN	90.2	84.2	80.9

Most of the comparatives representations presented in this work contain information sets of up to 16 participants, and for all of them the information appears to be grouped into classes, which suggest the existence of relationships between the signal behavior of emotional states.

Another important aspect that may be noticed is that even though the amount of the initial information is reduced, we did not get any significant penalty in the recognition rates and our identification rates are comparable to those presented by some of the most important researchers in the field, such as Dr Muruarapan. However, the lack of a standard methodology to perform an emotion recognition task is one of the main problems that we faced in the development of this research, since most works implement distinct approaches and even distinct data sources, resulting in a very wide range of experimental procedures and results, making comparison between them unfeasible. The proposals presented by Dr Muruarapan and Dr Scherer underpin the efforts of the community by supporting the importance of affective computing and its impact on technological developments. Also the development of affective computing techniques are becoming more viable with the development of a wide variety of portable devices which facilitate information gathering processes, as well as digital processing techniques.

Because we work with a group of specialists focused on the physical rehabilitation of high performance athletes, the implementation of this algorithm in a real-time platform, combined with a micro-expressions recognizing technique is contemplated as future work to serve as a support tool to monitor patient progress. This is important because experts suggest that some of the athletes endanger their physical integrity, whether by their competitive desire or by frustration. Also almost any modern device has capability to collect and process large amounts of information, and therefore the possibility of developing systems to recognize the emotional states of a user is of growing interest.

Acknowledgements

I would like to thank CONACYT for making this project possible and the Technological Institute of Tijuana for providing necessary technical support for the implementation of this project.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) received no financial support for the research, authorship, and/or publication of this article.

Notes

i. From the Latin *emovere*, meaning ‘moving’ , ‘displacing’.
ii. From the Latin *affectus*, meaning ‘to afflict’ , ‘to shake’ , ‘to touch’ .
iii. From the Latin *animus*, meaning ‘basic attitude’ or ‘governing spirit’ .
iv. Information processing (CNS), Support (CNS, NES, ANS), Executive (CNS), Action (SNS), Monitor (CNS).
v. (CNS), central nervous system; (NES), neuro-endocrine system; (ANS), autonomic nervous system; (SNS), somatic nervous system (the organismic subsystems are theoretically postulated functional units or networks).
vi. The motor area is considered as a reference for the filtering process, since this model considers the noise that is produced by the motor activity as eye movement or muscle activity.
vii. Since the records implemented in the experimental process were obtained from the DEAP dataset, it is expected that the emotional processes were heavily related to visual and auditory activity; however we also consider the limbic area as a region of interest, because most of the literature refers to it as emotion’s center in the brain, which it is also conveniently located in the temporal and parietal cortex areas.
viii. The same methodology could be applied to exclude the 29.5% of the electrodes of the 10/20 classical system and achieve a data reduction of 3,072 samples per second.
ix. It is important to note that both filters are interrelated, due to the fact that outputs of the transformation process deliver the detail (high pass filter) and approximation (low pass filter) coefficient that are implemented in the classification problem.
x. The same procedure was performed to acquire the approximation.
xi. The neural network inputs sorted the number of classes by manipulating the *a* parameter.
xii. This architecture was selected based on the experimental observations, where each of them were performed under a 10-fold cross-validation process.

References

1. Picard RW. *Affective computing*. 1st ed. Cambridge, MA: MIT Press, 2000.
2. James W. What is an emotion? *Mind* 1884; 9(34): 188–205.

3. Murugappan M, Rizon M, Nagarajan R, et al. Lifting scheme for human emotion recognition using EEG. In: *International symposium on information technology*, Kuala Lumpur, Malaysia, 26–28 August 2008, Vol. 2, pp.1–7. IEEE.
4. Murugappan M. Human emotion classification using wavelet transform and NN. In: *International conference on pattern analysis and intelligent robotics (ICPAIR)*, Putrajaya, 28–29 June 2011, Vol. 1, pp.148–153. IEEE.
5. Koelstra S, Muehl, Soleymani M, et al. DEAP: A database for emotion analysis using physiological signals,. *IEEE T Affective Comput* 2012; 3: 18–31.
6. Lin YP, Wang CH, Wu TL, et al. Multilayer perceptron for EEG signal classification during listening to emotional music. In: *TENCON 2007 – 2007 IEEE region 10 conference*, Taipei, 30 October–2 November 2007, pp.1–3. IEEE.
7. Heraz A, Razaki R and Frasson C. Using machine learning to predict learner emotional state from brainwaves. In: *Seventh IEEE international conference on advanced learning technologies*, Niigata, 18–20 July 2007, pp.853–857. IEEE.
8. Takahashi K. Remarks on SVM-based emotion recognition from multi-modal bio-potential signals. In: *13th IEEE international workshop on robot and human interactive communication*, 20–22 September 2004, pp.95–100. IEEE.
9. Li M and Lu BL. Emotion classification based on gamma-band EEG. In: *Annual international conference of the IEEE engineering in medicine and biology society*, Minneapolis, MN, 3–6 September 2009, pp.1223–1226. IEEE.
10. Rozgic V, Ananthakrishnan S, Saleem S, et al. Ensemble of SVM trees for multimodal emotion recognition. In: *Asia-Pacific signal information processing association annual summit and conference*, Hollywood, CA, 3–6 December 2012, pp.1–4. IEEE.
11. Attabi Y and Dumouchel P. Emotion recognition from speech: WOC-NN and class-interaction. In: *11th international conference on information science, signal processing and their applications*, 2–5 July 2012 pp.126–131. E-ISBN : 978-1-4673-0380-4 Print ISBN: 978-1-4673-0381-1INSPEC Accession Number: 13033317 Conference Location : Montreal, QC DOI: 10.1109/ISSPA.2012.6310487 Publisher: IEEE.
12. Razak A, Komiya R, Izani M, et al. Comparison between fuzzy and NN method for speech emotion recognition. In: *Third international conference on information technology and applications*, Montreal, QC, 4–7 July 2005, Vol. 1. pp.297–302. IEEE.
13. Gunler M and Tora H. Emotion classification using hidden layer outputs. In: *International symposium on innovations in intelligent systems and applications*, Trabzon, 2–4 July 2012, pp.1–4. IEEE.
14. Claparede E. Feelings and emotions. *Feelings and Emotions: The Wittenberg Symposium* 1928; 16: 124–139.
15. Cannon WB. The James– Lange theory of emotion: A critical examination and an alternative theory. *Am J Psych* 1927; 39: 10–124.
16. Asada M. Development of artificial empathy. *Neuroscience Res* 2015; 90: 41–50.
17. Calvo RA, D'Mello S, Gratch J, et al. *The Oxford Handbook of affective computing*. 1st ed. New York, NY: Oxford Library of Psychology, 2014.
18. do Amaral JR and de Oliveira JM. Limbic system: The center of emotions, http://www.healing-arts.org/n-r-limbic.htm (accessed January 2016).
19. Scherer KR. What are emotions? and how can they be measured? In: *Trends and developments: research on emotions*, 2005, Vol. 44, pp.695–729. Thousand Oaks, CA and New Delhi: Social Science Information and SAGE Publications.
20. Darwin C. *The expression of the emotions in man and animals*. 3rd ed. New York, NY: Oxford University Press, 1998.
21. Black J. Darwin in the world of emotions. *J Roy Soc Med* 2002; 95: 311–313.
22. Russell JA. A circumplex model of affect. *J Pers Soc Psychol* 1980; 39: 1161–1178.
23. Lehmann A, Bahçesular K, Brockmann EM, et al. Subjective experience of emotions and emotional empathy in paranoid schizophrenia. *Psych Res* 2014; 220: 825–833.
24. Roseman IJ. Phenomenology, behaviors, and goals differentiate discrete emotions. *J Pers Soc Psychol* 1994; 67: 206–221.

25. Roseman IJ. *Appraisal processes in emotion: Theory, methods, research*. 7th ed. Oxford: Oxford University Press, 2001.
26. Ekman P, Friesen WV, O'Sullivan M, et al. Universals and cultural differences in the judgments of facial expressions of emotion. In: *J Pers Soc Psychol* 1987 53: 712–717.
27. Ekman P and Friesen WV. Constants across cultures in the face and emotions. *J Pers Soc Psychol* 1971; 17: 124–129.
28. Clore GL and Ortony A. Psychological construction in the OCC model of emotion. *Emotion Rev* 2013; 5: 335–343.
29. The cognitive structure of emotions. Andrew Ortony, Gerald Clore and Alan Collins. New York, NY: University of Cambridge, 1988. no. of pages: 207. isbn 0-521-35364-5. *Applied Cognitive Psychology* 1992; 6: 181–182.
30. Lang PJ, Bradley MM and Cuthbert BN. International affective picture system (IAPS): Technical manual and affective ratings. In: *NIMH Center for the Study of Emotion and Attention, FL: The Center for Research in Psychophysiology, University of Florida*. Retrieved from http://csea.phhp.ufl.edu/Media.html#midmedia
31. Lang PJ, Bradley MM and Cuthbert BN. International affective digitized sounds (IADS): Stimuli, instruction manual and affective ratings. In: *The Center for Research in Psychophysiology, University of Florida*. Retrieved from http://csea.phhp.ufl.edu/Media.html#midmedia
32. Yin L, Wei X, Sun Y, et al. A 3D facial expression database for facial behavior research. In: *7th international conference on automatic face and gesture recognition*, Southampton, 2–6 April 2006, pp.211–216. IEEE.
33. Moody G, Mark R and Goldberger A. Physionet: A web-based resource for the study of physiologic signals. *IEEE Eng Med Biol Mag* 2001; 20: 70–75.
34. Petridis S, Martinez B and Pantic M. The {MAHNOB} laughter database. *Image Vision Comput* 2013; 31: 186–202.
35. Laurence J and Garey KB. *Brodmann's localisation in the cerebral cortex*. 1st ed. New York, NY: Spinger Science+Bussiness Media, 2006.
36. Jacobs KM. Brodmann's areas of the cortex. In: *Encyclopedia of clinical neuropsychology*. New York: Springer New York, 2011, pp.459–459.
37. Sanchez-Panchuelo RM, Besle J, Beckett A, et al. Within-digit functional parcellation of Brodmann areas of the human primary somatosensory cortex using functional magnetic resonance imaging at 7 tesla. *J Neurosci* 2012; 32: 15815–15822.
38. Isaacson RL. *Limbic system*. New York, NY: John Wiley and Sons, 2001.
39. LeDoux JE. Emotion circuits in the brain. *Ann Rev Neuroscience* 2003; 23: 154–184.
40. MacLEAN P. The limbic system (' visceral brain') and emotional behavior. *AMA Arch Neurol Psychiatry* 1955; 73: 130–134.
41. Sherwood L. *Human physiology from cells to system*. 8th ed. Boston, MA: Brooks/Cole, Cengage Learning, 2013.
42. Fox SI. *Human physiology*. 7th ed. New York, NY: McGraw-Hill, 2002.
43. Vidakovic B and Mueller P. Wavelets for kids: A tutorial introduction. Technical report, Duke University, 1991.

Health Informatics Journal

Article

Health Informatics Journal
2018, Vol. 24(2) 171–181

DOI: 10.1177/1460458216661863
journals.sagepub.com/home/jhi
SAGE

Effect of technology on aging perception

MA Rodrigo Juárez
Centro de Investigación Científica y de Educación Superior de Ensenada, Mexico

Víctor M González
Instituto Tecnológico Autónomo de México, Mexico

Jesús Favela
Centro de Investigación Científica y de Educación Superior de Ensenada, Mexico

Abstract
Technology can assist older adults to maintain an active lifestyle. To better understand the effect that technology has on aging perception, we conducted two studies. In the first study, through supraliminal priming, we analyzed the effects of aging- and technology-related stimuli on age estimation. In the second study, we conducted a technological intervention with a group of elders who used four interactive devices and analyzed effects on perceived aging. Results showed that technology-related stimuli did not affect estimated age. From the second study, we generated a sociotechnical model that explains the processes connecting technology use with successful aging. We concluded that the use of technology affects aging perception, although it depends on whether the elder people have a proactive attitude toward their aging process a priori.

Keywords
age estimation, aging perception, grounded theory, priming, technology

Introduction

Within the context of social informatics, a relevant phenomenon of study is the demographic transition characterized by the decline of birth and death rates[1] and its association with an increase in the usage of different technologies to enhance life quality and well-being. According to the United Nations' World Population Ageing 2013 report, the estimated global number of older adults (over

Corresponding author:
Víctor M González, Department of Computer Science, Instituto Tecnológico Autónomo de México, Río Hondo No. 1, Col. Progreso-Tizapán, 01080 Mexico City, Mexico.
Email: victor.gonzalez@itam.mx

60 years) for 2013 was 841 million people. This number is expected to increase to 2 billion people in 2050; this will represent 21.1 percent of the global population.[2] Regarding the second element, mobile phone and Internet usage are growing rapidly. Taking Mexico as an example, it has been estimated that the number of older adults who could access Internet services from 2007 to 2015 increased from 2 percent (www.amipci.org.mx/estudios/habitos_de_internet/2007_Habitos_Usuarios_Internet_Mx-1.pdf) to 4 percent (www.amipci.org.mx/images/amipci_habitos_del_internauta_mexicano_2015.pdf) of the total Internet users. Around the world and in places with good access to information and communication services and networks, the presence of technology moves quickly beyond the desktop and starts creating collections of interconnected devices that become part of the entire space where people live. For many people, and in particular for older adults, these technologies can become ambient-assisted living solutions which can help them to improve their health and quality of lives.[2]

These forces changing demographics and the pervasiveness of technology would, at the end, result on changes in the way people will experience aging and, in general, their perception of aging. Technology usage by elder people is increasing as it incorporates its needs and worries,[3,4] even though negative stereotypes regarding older adults and technology prevail in society.[5,6] The fact that designers and developers often fail to consider older adults' needs is a factor that might contribute to create scenarios where older adults look incompetent when dealing with new technologies.[4,7,8]

Aging perception is a multidimensional construct that plays a significant role in society since it triggers different attitudes and behaviors toward the older adults and it affects their health and well-being. For example, the dimensions of physical appearance and aging self-perceptions provide significant cues about health, lifestyle or longevity that help us to allocate the person within an age range, therefore affecting the way we talk and behave toward them, and are associated with different elements such as stereotypes and life experiences.[9–12] Different studies have shown that older adults with positive "Aging Self-Perception" (ASP) tend to have lower rates of suicide and depression,[13] heart diseases,[14] hypertension disorders[15] and general mortality,[16,17] which translate into longer and more joyful lives. Consequently, we need to understand the different dimensions of the relationship between older adults and technology starting from basic aspects such as people becoming used to the idea of technology being used by older adults.

The main purpose of this research is to show the effects that technology produces in two dimensions of aging perception, age estimation and self-perception. This is important for technology-based ambient-assisted living because it would mean that technology could provide a context and a set of tools to improve, on one hand, behaviors of younger people toward older adults (which could mean an enhanced interaction and inclusion to the digital world) and, on the other hand, their own behaviors in order to achieve a successful and healthy aging.

Previous work

Technology has the potential of modifying the way we perceive the world. This encompasses social processes such as the assimilation of a new culture,[18] perceptions toward people with conditions that affect their behavior (e.g. autism),[19] intergenerational interactions, decreasing the stress due to lifestyle changes (e.g. new independence)[20] and life transitions.[21]

Several studies have considered how different factors influence age perception. These include subjective and objective age, attractiveness and facial aging cues[12] and stereotypes.[22,23] Although scarce, there is prior research that shows that technological tools that promote shared activities can modify people's perception on aging. Particularly, by creating contexts that promote a more symmetric interaction between young people and elders, thus discrediting some stereotypes related to technology.[24,25]

Thus, even though there have been numerous studies on how different factors can alter the perception of aging, the use of technology has not been studied as being one of them. Therefore, this article tries to fill this void and generate a better understanding of this relationship. To achieve this, we divided our investigation in two studies. In the first one, we addressed the effect of technology- and aging-related stimuli on age estimation. In the second study, we tried to answer how does technology affects the aging self-perception in a group of older adults and explored a set of technological paradigms as they were used by elderly participants for a number of weeks.

First study

The focus of this study was to determine if technology has an indirect effect on age estimation through supraliminal priming. We formulated two hypotheses to be tested: (1) participants exposed to stimuli related to technology will assign lower age ranges to people depicted in different photographs than those exposed to neutral stimuli and (2) participants exposed to stimuli related to aging will assign higher age ranges to people depicted in different photographs than those exposed to neutral stimuli.

Methodology

The methodology of this first study is fully reported in Juarez;[26] therefore, we will mention it briefly, focusing on its results and how it connects to the next study. We recruited 60 participants (22 men and 38 women, aged 19–88 years) from the student and faculty bodies of a research center in Mexico and from older adults participating in a course. The participants were randomly assigned to one of three groups (technology, aging and neutral) according to the stimulus to which they would be exposed. Each participant had an individual session in which they completed five activities: a priming activity, a decoy activity, a photography test, a debriefing activity and a personality questionnaire. The priming activity consisted of the "Sentence Unscrambling Task,"[27] there was no time limit for this activity in order to guarantee supraliminal priming. To reduce the possibility of recalling the priming words in the photography test, we extended the time between the priming and the real test by implementing a 10-min decoy activity. For the photography test, each participant had to assign an age range to different people depicted in a set of photographs. As exclusion criteria, we measured the degree of awareness about the stimuli through a funnel debriefing questionnaire[28] and the tendency to have sudden mood changes through the "Neuroticism" subscale of the short version of the revised Eysenck Personality Questionnaire Brief Version (EPQ-BV) for adults.[29]

Results

Five participants had a score higher than 6 in the "Neuroticism" subscale; therefore, their results were overlooked in the subsequent analysis. Thus, the total number of participants was 55 (18 for the "Neutral" group, 19 for the "Technology" group and 18 for the "Aging" group). Six of the participants noticed that the words used in the priming were related, although none of them realized the association between the priming and the objective or between the priming and the photography test; therefore, according to Aronson et al.,[30] none of the remaining participants were excluded. The analysis of the photograph test comprised a chi-square association test among the mode of the age ranges estimated by the participants of each group. Results showed that in 53 percent of the photographs, the mode of the age ranges was the same among the different groups. Although in 47 percent of the pictures the estimated age was different, the statistical test showed that this was not statistically significant ($p < 0.05$, $\alpha = 5\%$).

Discussion

After the analysis, we rejected both of our hypotheses, since the estimated age ranges from the participants of both experimental conditions did not show statistical significant difference from the estimations of the neutral condition. Considering the conclusions from Ostir et al.[15] and Eibach et al.,[23] the results were unexpected, since they found that aging-related stimuli could induce participants to perceive other people as older. The lack of statistical significance from the technology- and aging-related stimuli may be due to the following: technology, by itself, not being a factor that alters the perception of aging, at least unconsciously, prejudices toward technology or cultural factors. Despite these results, previous research[24] shows that it is possible for technology to affect attitudes and perceptions toward aging due to its impact in some of its dimensions such as emotions, learning, satisfaction with life and promoting physical and mental activity. In order to explain the mismatch between our results and previous research, we needed to address the problem from another point of view.

Particularly, among older adults, the effects of prejudices toward technology are more important, since its pervasiveness means that more aspects of their life are now influenced in a negative way. Therefore, the adoption and usage of technology can be used to mitigate or eliminate these prejudices in order to explore how older adults change their aging perception, as a whole, through technology. If technology can have a positive effect on the perception of aging, then technological infrastructure such as ambient computing to support assisted living could be used as a mechanism to promote "active aging" and to increase quality of life. This could produce positive effects such as improving physical and mental health, enhancing communication and socialization, disposing negative stereotypes about older adults and technology and helping them maintaining their independence.

These results lead us to design a new study focusing on the influence that direct contact with technology has on the way older adults perceive and live their aging process.

Second study

Methodology

The objective of this study was to understand how frequent use of technology, even for just a few weeks, could affect the self-perceptions and experiences toward aging. At this point, our investigation aimed to understand the phenomenon from a more open and naturalistic perspective where a number of contextual factors in the lives of the older adult can be considered. Given the nature of the study, we decided to use qualitative methods of data collection and analysis.

Recruiting strategy. Using the purposeful sampling method,[31–33] we recruited six older adults (five women and one man) whose ages ranged between 65 and 84 years ($\mu = 76.14$ years, $\sigma = 7.13$ years). Three of them attended a basic informatics course, and the rest of them attended to different non-technology activity groups (e.g. cooking), and all of them were mobile phone users. We looked for these two features due to the effects of different technological expertise while coping with new technologies[34–36] without resulting any practical advantage[37,38] and to ease the adoption of new technologies, as our focus was the use of technology rather than the process of adoption.

Data collection and analysis. Each participant had to use four technological devices representing different technological paradigms (a mobile phone, an online social network, an e-reader and an activity tracker). In order to augment the likelihood that perception change was due to the use of technology, for every device, a series of tasks were defined so that we could monitor the activity of the participants during the period of usage. We conducted four semi-structured interviews with

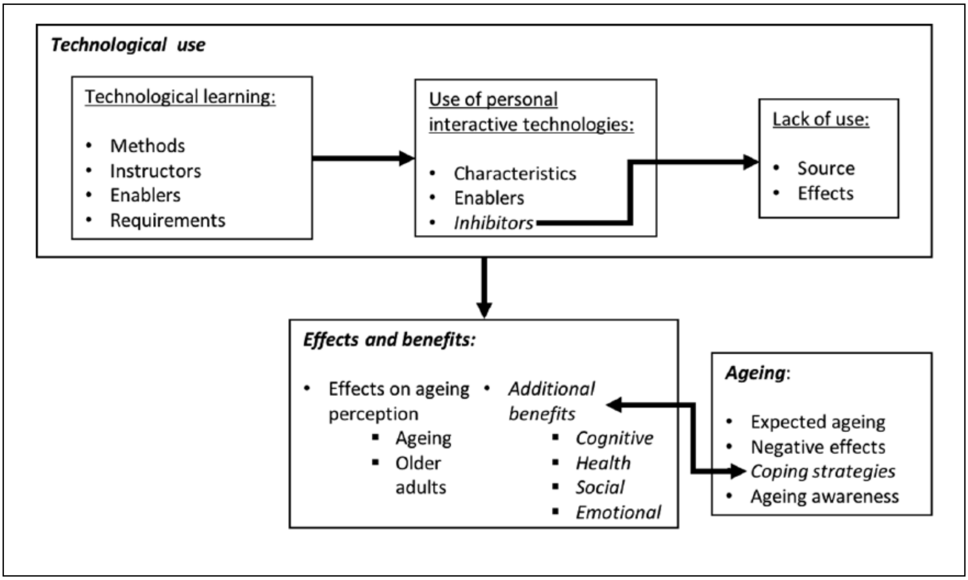

Figure 1. A model of technological use its effects and influence on aging perception.

each of the participants (approximately every 2 weeks) at their home or in a waiting room of a social center. We opted this kind of interviews due to the possibility of exploring emerging topics regardless of the possibility of different questions to the participants.[39,40] In the first session, the interviews focused on their personal experiences with older adults, ASP, technological background, conceptualization of technology, technology and aging association and subjective age. The following sessions focused on the device usage, conceptualization of technology, technology and aging association and subjective age. Once the interview was over, we showed them how to use the technological device they were going to use for the next couple of weeks.

To analyze the interviews, we used a qualitative approach based on the grounded theory methodology.[41] The interviews were transcribed and analyzed using this methodology, and through constant comparison, we produce a model explaining the phenomenon. This theoretical model is compared to existing related frameworks in order to detect similarities and differences, thus increasing the internal validity and the generality of the new theoretical model.[42]

Results

From the data, a model emerged which organizes the findings around three main categories showing that *technological use* produces a series of *effects and benefits* that influence the process of *aging* and its perception (Figure 1). In particular, there is a set of cognitive, social, emotional and health-related benefits that match the effects of different coping strategies that the older adults implement to improve their well-being and achieve a successful aging.

Technological use. Older adults need a learning process in order to use new technological devices, as one participant exemplifies,

> … I mean, if you tell me "This is how it is done," then maybe I'll practice it and learn how to do it. (Participant 2)

This process can be carried out in different ways, such as attending formal (courses) and informal classes (with relatives or other older adults) or exploring the new technology by themselves. Attending to formal courses exposes the older adults to others; this promotes new interactions and a constant exchange of ideas and experiences to deal with new technologies. One of the most important requirements for this process is the wish to learn new skills; it promotes commitment to get involved in different activities to learn how to use new technologies and helps them remember past periods of their life, for example, their student years, as one participant said,

> I go to my computer class and I feel young. I grab my backpack, take my bus (laughs) just like any student, raining or not. (Participant 6)

As a result of this learning process, the older adults are able to use new technological devices. Although it is important to notice that this does not represent an end of the learning process, as it can occur while they are learning or as a promoter to learn other technologies.

Their use of technology is characterized by their role as consumer and producer of digital information, additional tools to support them (e.g. user manuals) and the different objectives they pursue while using them. Anonymity is an important enabler because it lets the older adults interact in a virtual context without being subject to ageist attitudes and stereotypes. Environment also plays a significant role in the use of technology as it can work as an inhibitor or an enabler. For example, one participant said that her lack of interest in technology was due to her non-technological environment. This had a significant impact because it hindered the learning and usage process.

Two of the technological devices' characteristics that inhibit their usage are the difficulty to use and the diffusion of digital contents. When difficulty to use and a limited knowledge (perceived or real) interact, older adults tend to feel uneasy about the technology highlighting their fears to commit errors that could damage the device, as one participant said. Although these fears can decrease, if they own the technology they are using. Excessive usage can also have a negative effect on the attitudes toward technology, as they think that using devices in inappropriate situations can damage the relationship with others. These can produce a lack of use that has important effects, given society's technological influence. The participants perceived that the older adults who do not use technology get isolated from other people and are left behind in the evolution of society.

Effects and benefits. We found that the use of technological devices produces a series of effects in different areas of self-perception by older adults. The following two quotes illustrate the effects on the perception of an older adult and aging as concepts characterized by a constant effort to keep up with the constant changes in technology and society in order to be connected and included:

> They (grandchildren) told me that I was very modern (laughs) … (Participant 1)

> … I told him (husband) that he was old. I think that my husband stayed behind. I think that … I'm more to date with society. (Participant 4)

But one of the most important effects concerns to their role in society. On one hand, these older adults can be a role model for younger generations, showing them that aging is not necessarily a negative stage of life and they need to embrace the new technologies and challenges. On the other hand, they assume the role of being a technology enabler for other older adults, which gives them a prominent position in their communities, as one participant pointed out. Additionally, there are a set of benefits in other areas such as cognition, socialization emotions and health:

> I feel that I can teach them (other older adults), at least, this (using the FitBit) … I've been talking about it and I have noticed a lot of positive reception (sic) on their behalf. Now that this (FitBit) is a novelty I can motivate them to do more exercise … (Participant 1)

Among the cognitive benefits emerged learning something new as an activity to keep using the brain and thus preventing its decay, the access to lots of information about topics of their interest (e.g. news, diseases or new treatments) and a mental challenge to use new technologies.

Using technology also makes them feel younger, active, independent, happier and more satisfied with their lives, and important within their communities, improving their quality of life and well-being.

Given the fact that older adults' social circles tend to narrow due to illness or death, it is very important for them to create new relationships and to maintain the existing ones. Therefore, technology is considered an important tool for them, as it keeps them included in their relatives' and friends' lives. It also helps them to communicate with people they are not acquainted with but who share common interests with them. This helps them reduce the feeling of isolation and abandonment.

We noticed that participants became aware of the need to maintain physically and mentally active, therefore promoting behavior change, as it motivates them to find new activities or retake activities that were abandoned or look for information about their conditions. Although, the use of technology represents a new challenge as they need to maintain those healthy behaviors even without the presence of the technological aids:

> … I stopped napping, now I walk and one very important thing is that I quit smoking … I restarted walking- Before, about 6 months ago, I jogged but I stopped doing it. (Participant 6)

Aging. The aging perception forms since early life stages. In most cases, it comes from the experiences that people had with their relatives (e.g. grandparents). These experiences made them aware of the negative physical and lifestyle changes that aging involves, which become a point of comparison with their own aging. When we asked a participant about the aging process of one of her parents, she said,

> … my dad didn't go out anymore because he had an injury in his foot. Then, he locked himself in his world … (Participant 2)

Then, when asked about how that event had influenced her, she said,

> … I'm not going to lock myself until I can't, really, really look out for myself. (Participant 2)

In order to cope with these changes, older adults implement different strategies. These strategies have diverse objectives such as accepting the physical and emotional changes, focusing on the positive aspects of aging, keeping an acceptable activity level and socializing and compensating their losses. These strategies require from the older adult willingness and an open attitude to discover and engage in new activities. As a result of the implementation of these strategies, the older adults achieve a successful and healthy aging. This kind of aging is characterized by factors such as independence, dignity and energy and is perceived as a stage of personal growth.

Discussion

The last phase of the grounded theory methodology involves a comparison between the model developed and related models that exist in the literature. In order to make this comparison, we divided our model in two parts: a technological part (that includes the learning and use of technology process) and the aging part.

The technological part of the model was compared with the diffusion of innovations model (DIM)[43] and the technology acceptance model (TAM).[44] The main differences with DIM[43] are (1) that our model only comprises the process of use of technology rather than the adoption process and (2) that our model considers that the use of technology rather than the technology itself should adjust the social norms and etiquette. Regarding TAM,[44] the main similarity is that the perceived ease of use is not a determinant factor while deciding to use or not a device. This is due to their awareness of their lack of knowledge of new technologies; therefore, difficulties while using it are expected and can be overcome with time and practice.

The aging part of the model was compared with the model of successful aging of elders with late-life disabilities (MSAELD)[45] and the healthy aging model (HAM).[46] The main difference with MSAELD[45] is that our model highlights the awareness of the aging process rather than ignoring it. While with HAM,[46] our model concurs with the effect of the social resources to improve aging, although our model focuses on the technological elements that enhance the communication with those social resources.

One of our main results showed that technology should not be seen as a tool that by itself is capable of producing positive outcomes in older adults' health and lifestyle. It is a tool that should be coupled with coping strategies in order to amplify and diversify them. This entails that the older adults must adopt an active role in their aging process that drives them to implement the above said strategies, which confirms the results obtained in Salovaara et al.[21] Also, we found that the usage of common technological tools create an anonymous context that promotes intergenerational communication (even with strangers), thus making older adults feel included in their families, and thus generating and consolidating new identities away from social stigmatization; this is in accordance with previous works.[18–21,24,25] Despite these benefits, older adults showed different issues related to social norms on digital environments, as they tried to transfer their "real-world" social norms.

In general, our results showed that the use of technology produces different benefits that can derive in a reduction in health-related problems. Among those benefits are greater socialization that is related to fewer depressive events and suicides, the change to healthy habits (such as exercise and better diets) that is related to fewer heart- and pressure-related conditions and, overall, an increased emotional and physical well-being that is related to lower indexes of general mortality.

Finally, our results showed that the lack of usage increases the feeling of isolation and abandonment as people are left behind in a society that constantly evolves.

Conclusion

From the second study, we can conclude that the use of technology affects aging perception disrupting preconceived ideas on the subject (which are heavily negative related), although this effect is not direct. The results showed that these effects occur through the correspondence between the coping strategies that the older adult implements and the benefits they gain from the usage of the technological devices (Figure 2) (such as keeping social circles and accepting physical and mental changes). However, these positive effects depend on different intrinsic factors such as the older adult's involvement in changing their aging experience and perception, their socioeconomic context and the coping strategies they implement.

It is important to consider that some of the benefits that emerged may be specific to the kind of technological devices that were chosen, and this may entail that other devices may have other results. Due to time constraints, we focused on the immediate usage and effects of technology rather than in its adoption and long-term effects. Finally, our results showed that the usage of

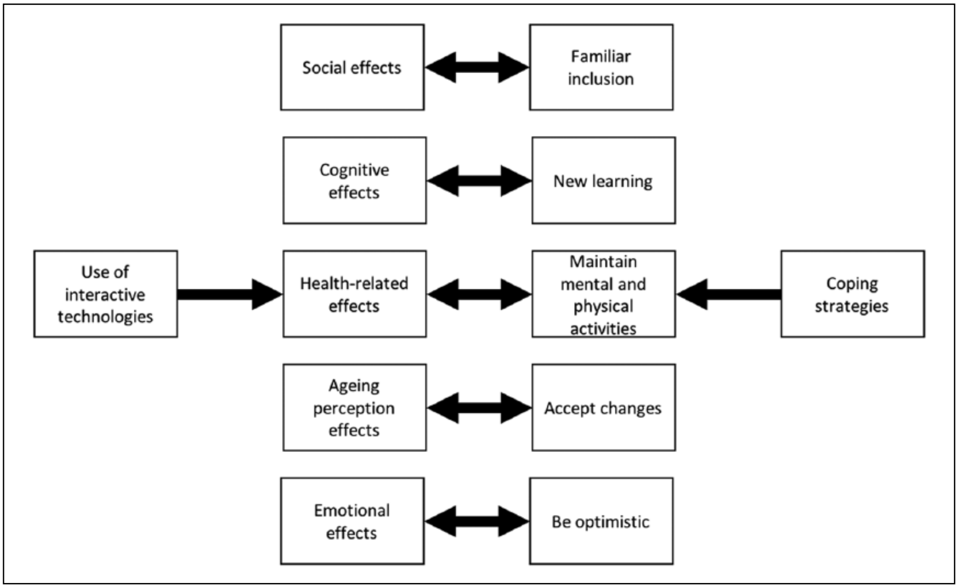

Figure 2. Relationship between use of interactive technologies and coping strategies.

technology produced behavior changes that may lead to a healthier lifestyle (e.g. health awareness, resume physical activity and socialization), although it is necessary to conduct a longitudinal approach that provides conclusive data about these changes and their long-term health effects.

Acknowledgements

The authors want to thank all the participants for their help and enthusiasm during both studies, CICESE and UABC for their help during the recruitment and for letting us use their facilities, ITAM for their help providing the equipment for the second study, INAPAM and the San Francisco Social Center for their support during the recruitment and interviews and the Mexican Culture Association A.C. for their support.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This work was supported by the Mexican Science and Technology National Council (CONACyT) (grant number 367827).

References

1. Centre for Health Development. A glossary of terms for community health care and services for older persons, http://www.who.int/kobe_centre/ageing/ahp_vol5_glossary.pdf (2004, accessed 19 August 2014).
2. Thomas K, Becker M, Ras E, et al. Ambient intelligence in assisted living: enable elderly people to handle future interfaces. In: *4th international conference on universal access in human-computer interaction, UAHCI 2007 Held as part of HCI international*, Beijing, China, 22–27 July 2007, pp. 103–112. Berlin: Springer.

3. Hope A, Schwaba T and Piper AM. Understanding digital and material social communications for older adults. In: *Proceedings of the 32nd annual ACM conference on human factors in computing systems—CHI '14*, Toronto, ON, Canada, 26 April–1 May 2014, pp. 3903–3912. New York: ACM.
4. Östlund B. Silver age innovators: a new approach to old users. In: Kohlbacher F and Herstatt C (eds) *The silver market phenomenon*. Berlin; Heidelberg: Springer, pp. 15–26.
5. Barrett AE and Pai M. Sketches in cyberspace: using student drawings of elders in an online social gerontology course. *Gerontol Geriatr Educ* 2008; 29: 84–103.
6. Hernández-Encuentra E, Pousada M and Gómez-Zúñiga B. ICT and older people: beyond usability. *Educ Gerontol* 2009; 35: 226–245.
7. Blythe MA, Monk AF and Doughty K. Socially dependable design: the challenge of ageing populations for HCI. *Interact Comput* 2005; 17: 672–689.
8. Sun Y, Ding X, Lindtner S, et al. Being senior and ICT. In: *Proceedings of the 32nd annual ACM conference on human factors in computing systems—CHI '14*. Toronto, ON, Canada, 26 April–1 May 2014, pp. 3933–3942. New York: ACM.
9. Anzures G, Ge L, Wang Z, et al. Culture shapes efficiency of facial age judgments. *PLoS ONE* 2010; 5: e11679.
10. National Centre for the Protection of Older People. Public perceptions of older people and ageing: a literature review, http://www.ncpop.ie/userfiles/file/ncpopreports/Review1LROlderpeopleandageing.pdf (2009, accessed 15 August 2014).
11. Levy BR. Mind matters: cognitive and physical effects of aging self-stereotypes. *J Gerontol B-Psychol* 2003; 58: 203–211.
12. Kotter-Gruhn D and Hess TM. So you think you look young? Matching older adults' subjective ages with age estimations provided by younger, middle-aged, and older adults. *Int J Behav Dev* 2012; 36: 468–475.
13. Koivumaa-Honkanen H, Honkanen R, Viinamäki H, et al. Life satisfaction and suicide: a 20-year follow-up study. *Am J Psychiatry* 2001; 158: 433–439.
14. Davidson KW, Mostofsky E and Whang W. Don't worry, be happy: positive affect and reduced 10-year incident coronary heart disease: the Canadian Nova Scotia Health Survey. *Eur Heart J* 2010; 31: 1065–1070.
15. Ostir GV, Berges IM, Markides KS, et al. Hypertension in older adults and the role of positive emotions. *Psychosom Med* 2006; 68: 727–733.
16. Xu J. *Subjective well -being as predictor of mortality, heart disease, and obesity: prospective evidence from the Alameda County Study*. Texas Medical Center Dissertations, (via ProQuest), Texas Medical Center, Houston, TX, pp. 1–125, 2005.
17. Keyes CLM and Simoes EJ. To flourish or not: positive mental health and all-cause mortality. *Am J Public Health* 2012; 102: 2164–2172.
18. Zhang S, Jiang H and Carroll JM. Babel or great wall: social media use among Chinese students in the United States. In: *Proceedings of the 30th ACM international conference on design of communication—SIGDOC'12*, Seattle, WA, USA, 3 October 2012, p. 37. New York: ACM.
19. Burke M, Kraut R and Williams D. Social use of computer-mediated communication by adults on the autism spectrum. In: *Proceedings of the 2010 ACM conference on computer supported cooperative work—CSCW '10*, Savannah, GA, 6–10 February 2010, p. 425. New York: ACM.
20. Smith ME, Nguyen DT, Lai C, et al. Going to college and staying connected: communication between college freshmen and their parents. In: *Proceedings of the ACM 2012 conference on computer supported cooperative work—CSCW '12*, Seattle, WA, USA, 11 February 2012, p. 789. New York: ACM.
21. Salovaara A, Lehmuskallio A, Hedman L, et al. Information technologies and transitions in the lives of 55–65-year-olds: the case of colliding life interests. *Int J Hum Comput Stud* 2010; 68: 803–821.
22. Kotter-Grühn D and Hess TM. The impact of age stereotypes on self-perceptions of aging across the adult lifespan. *J Gerontol B-Psychol* 2012; 67: 563–571.
23. Eibach RP, Mock SE and Courtney EA. Having a "senior moment": induced aging phenomenology, subjective age, and susceptibility to ageist stereotypes. *J Exp Soc Psychol* 2010; 46: 643–649.
24. Chua P-H, Jung Y, Lwin MO, et al. Let's play together: effects of video-game play on intergenerational perceptions among youth and elderly participants. *Comput Hum Behav* 2013; 29: 2303–2311.

25. Cornejo R, Tentori M and Favela J. Enriching in-person encounters through social media: a study on family connectedness for the elderly. *Int J Hum Comput Stud* 2013; 71: 889–899.
26. Juarez MAR, Favela J and Gonzalez VM. Do technology-related stimuli affect age estimation? In: Cleland I, Guerrero L and Bravo J (eds) *Ambient assisted living: ICT-based solutions in real life situations*. Cham: Springer International Publishing, 2015, pp. 259–264.
27. Srull TK and Wyer RS. The role of category accessibility in the interpretation of information about persons: some determinants and implications. *J Pers Soc Psychol* 1979; 37: 1660–1672.
28. Bargh JA and Chartrand TL. The mind in the middle: a practical guide to priming and automaticity research. In: Reis HT and Judd CM (eds) *Handbook of research methods in social and personality psychology*. Cambridge: Cambridge University Press, 2000, pp. 311–344.
29. Sato T. The Eysenck Personality Questionnaire Brief Version: factor structure and reliability. *J Psychol* 2005; 139: 545–552.
30. Aronson E, Wilson TD and Brewer MB. Experimentation in social psychology. In: Gilbert DT, Fiske ST and Lindzey G (eds) *The handbook of social psychology*. New York: McGraw-Hill, 1998, pp. 99–142.
31. Creswell JW and Plano Clark V. *Designing and conducting mixed methods research*. 2nd ed. London: SAGE Publications Inc, https://us.sagepub.com/en-us/nam/designing-and-conducting-mixed-methods-research/book233508 (2011, accessed 15 July 2015).
32. Palinkas LA, Horwitz SM, Green CA, et al. Purposeful sampling for qualitative data collection and analysis in mixed method implementation research. *Adm Policy Ment Health*, 42(5): 533–544.
33. Bernard HR. Research methods in anthropology: qualitative and quantitative approaches, https://books.google.com/books?hl=es&lr=&id=WhKYqATAySwC&pgis=1 (2011, accessed 15 July 2015).
34. Engelbrektsson P, Yesil Ö and Karlsson ICM. Eliciting customer requirements in focus group interviews: can efficiency be increased? In: *EIASM, The 7th international product developement management conference*, http://www.researchgate.net/publication/252532750_ELICITING_CUSTOMER_REQUIREMENTS_IN_FOCUS_GROUP_INTERVIEWS_CAN_EFFICIENCY_BE_INCREASED (2000, accessed 15 July 2015).
35. Liu Y, Osvalder A-L and Karlsson M. Considering the importance of user profiles in interface design. In: Matrai R (ed.) *User interfaces*. Vukovar: InTech Publishing, 2010, pp. 61–80.
36. Liu Y and Osvalder A-L. Usability tests as a benchmarking tool—a case study on complex medical ventilators. In: Bust PD (ed.) *Contemporary ergonomics 2009*. London: CRC Press, 2009, pp. 182–188.
37. Juarez R and Gonzalez VM. Mental models, performance and usability of a complex interactive system: the case of Twitter. In: *2013 Mexican international conference on computer Science*, Morelia, Mexico, 30 October–1 November 2013, pp. 7–12. New York: IEEE.
38. Liu Y. Usability evaluation of medical technology: investigating the effect of user background and users' expertise, http://publications.lib.chalmers.se/publication/92255 (2009, accessed 14 July 2015).
39. Galletta A. Mastering the semi-structured interview and beyond: from research design to analysis and publication, https://books.google.com/books?hl=es&lr=&id=NdbtHg6sPgIC&pgis=1 (2013, accessed 10 July 2015).
40. Newton N. The use of semi-structured interviews in qualitative research: strengths and weaknesses, http://www.academia.edu/1561689/The_use_of_semi-structured_interviews_in_qualitative_research_strengths_and_weaknesses (2010, accessed 10 July 2015).
41. Glaser BG and Strauss AL. *The discovery of grounded theory: strategies for qualitative research*. London: Transaction Publishers, https://books.google.com/books?id=rtiNK68Xt08C&pgis=1 (2009, accessed 14 July 2015).
42. Eisenhardt KM. Building theories from case study research. *Acad Manage Rev* 1989; 14: 532–550.
43. Rogers EM. *Diffusion of innovations*. 5th ed. New York: Free Press, 2003.
44. Davis FD. User acceptance of information technology: system characteristics, user perceptions and behavioral impacts. *Int J Man Mach Stud* 1993; 38: 475–487.
45. Romo RD, Wallhagen MI, Yourman L, et al. Perceptions of successful aging among diverse elders with late-life disability. *Gerontologist* 2013; 53: 939–949.
46. Bryant LL, Corbett KK and Kutner JS. In their own words: a model of healthy aging. *Soc Sci Med* 2001; 53: 927–941.

Health Informatics Journal

Article

Health Informatics Journal
2018, Vol. 24(2) 182–193

DOI: 10.1177/1460458216661864
journals.sagepub.com/home/jhi
SAGE

Assessing empathy and managing emotions through interactions with an affective avatar

Esperanza Johnson, Ramón Hervás and Carlos Gutiérrez López de la Franca
University of Castilla-La Mancha, Spain

Tania Mondéjar
University of Castilla-La Mancha, Spain; eSmile, Spain

Sergio F Ochoa
Universidad de Chile, Chile

Jesús Favela
Centro de Investigación Científica y de Educación Superior de Ensenada (CICESE), Mexico

Abstract

Assistive technologies can improve the quality of life of people diagnosed with different forms of social communication disorders. We report on the design and evaluation of an affective avatar aimed at engaging the user in a social interaction with the purpose of assisting in communication therapies. A human–avatar taxonomy is proposed to assist the design of affective avatars aimed at addressing social communication disorder. The avatar was evaluated with 30 subjects to assess how effectively it conveys the desired emotion and elicits empathy from the user. Results provide evidence that users become used to the avatar after a number of interactions, and they perceive the defined behavior as being logical. The users' interactions with the avatar entail affective reactions, including the mimic emotions that users felt, and establish a preliminary ground truth about prototypic empathic interactions with avatars that is being used to train learning algorithms to support social communication disorder evaluation.

Keywords

affective computing, cognitive disabilities, empathy, human–avatar interaction, social communication disorder

Corresponding author:
Ramón Hervás, MAmI Research Lab, University of Castilla-La Mancha, MAmI Research Lab, Paseo de la Universidad 4, 13072, Ciudad Real, Spain.
Email: ramon.hlucas@uclm.es

Introduction

Social communication disorder (SCD) typically encompasses problems with social interaction, social understanding, and pragmatics.[1] Assistive technologies have the potential to assist individuals who suffer from the ailment to live more independent lives, facing less difficulties and socialization issues.[2] In particular, affective avatars can engage users in conversations that can lead to effective therapies that contribute to improve their social skills, the management of their emotions and empathy, and the understanding of themselves (introspection). In this article, we report on the design and evaluation of an affective avatar that interacts with people. The avatar evaluation involved interactions with normotypical users, who tried to interpret the emotion exhibited by the avatar and whether its behavior seemed like a logical response to the interaction.

SCDs entail multiple kinds of issues, and each of them can be faced through different treatments. Therefore, we defined a taxonomy of human–avatar interactions to support seven cognitive processes. Then, we used the taxonomy to instantiate this prototype, by adapting it to support a particular cognitive process related to communication: emotional states and empathic behavior, a singular area of intervention in SCDs.

Related work

Human–avatar interaction has been the subject of research for several years. It has been applied to assist older adults in conducting activities of daily living (ADLs),[3] assisting people with severe motor disabilities,[4] and supporting elderly people with mental disorders and/or physical disabilities.[5] Additional uses include a three-dimensional (3D) avatar as an assistant in interactive TV[6] and for engagement in videogames.[7] These previous works have helped us in the process of understanding what possible needs and technologies have been used in terms of assistance to people with or without disabilities.

Some studies have assessed user preferences with different types of avatars. Some were directed at specific populations such as older adults[3] or children.[8] In some cases, questionnaires are used to conduct these evaluations, and the avatar is not necessarily fully functional; a few images of the avatar might be sufficient.[8,9] Some of the findings of these studies include that the design of avatars needs to consider the needs and interests of children[8] and that a human-like appearance and a friendly attitude were preferred by older adults. A recent study explored how older adults perceived the emotion expressed in the facial expression of virtual avatars in three emotional dimensions—Evaluation (valance), Potency (power/dominance), and Activity—with significant consensus reached in the first two dimensions.[10] Of those avatars implemented, natural interaction has been achieved using gestures[5] as well as verbal communication.[7] Other studies have explored how to enact intelligent reactive behavior in avatars.[11] With regard to avatar, realism studies have found that visual realism associated with low kinematic conformity scored low in perception of social presence,[12] while Kang and Watt[13] found evidence that satisfaction with communication was associated with high anthropomorphic avatars. All these studies have helped shape the design of the avatar, as well as an awareness of what has been previously done with avatars in general.

When comparing with our work, previous research has placed no special emphasis on the affective reaction of the avatar and the emphatic response of the user. More importantly, we are not aware of previous research on the use of avatars to assist in dealing with communication disorders, generally having seen proposals with very specific goals of guiding people (either elderly or disabled) with certain activities, centered on improving the quality of life through simplifying daily tasks. The main novelty of this work is the quantitative analysis of interactions of neurotypical people with affective avatars that state the ground truth to develop future assistive avatars.

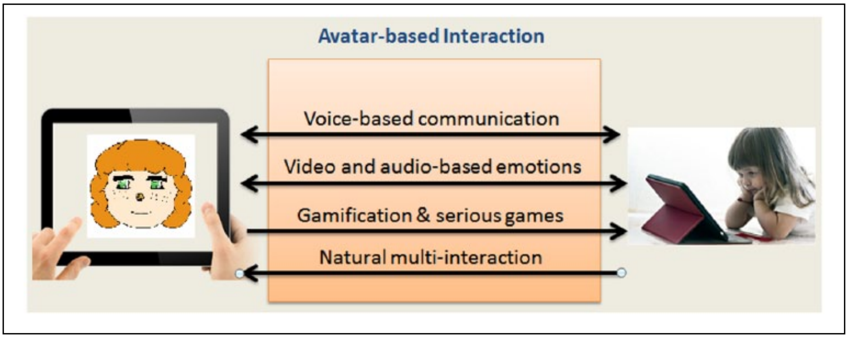

Figure 1. Avatar-based interactions included in the system and draft of the avatar.

Design of the avatar-based interactions

In Ref. 14, the authors presented a conceptual prototype of the avatar designed to interact with SCD patients. Figure 1 shows the interactions supported by the avatar, indicated by arrows, which point both ways indicating bilateral communication between the human and the avatar. Particularly, voice-based communication and video- and audio-based emotion interactions refer to the audio, verbal, ocular, and facial interactions that will be explained in the taxonomy presented below.

Gamification and serious games[15,16] refer to the narrative role playing games that the avatar will play with the human, although in this case it is unilateral as it is something that the avatar sends to the human (as the user does not propose games to the avatar). Natural multi-interaction comes from the human and they are the user's reactions to emotions, messages, and games that the avatar shows. So although it seems like the multi-interaction is unilateral, the whole interaction is bilateral between the human and the avatar, but they perform their interaction in different ways and with different understandings; therefore, the terminology changes when the avatar interacts with the human and when the human interacts with the avatar.

The interactions were implemented through gestures in a tablet-PC; therefore, the peripherals used as sensors are those available in this device type: a camera for ocular and facial interaction, microphone for verbal interaction, and headphones/speakers for audio interaction. Both the accelerometer and the tactile screen supported the tactile interaction.

Taxonomy of human–avatar interactions

The interactions that a person can have with an avatar can be classified into different categories. Each interaction type can be used to treat particular communicative and cognitive skills. For instance, treatments to improve communication skills in children with SCD involve verbal and non-verbal communication skills.

In order to help developers identify which interaction type can be used to enhance and improve particular communication skills, we defined a taxonomy that takes into account seven cognitive processes (Table 1): *joint attention* (*J*) is the shared focus of two individuals on a same point or object and also the action of looking at each other's eyes; *focused attention* (*F*) is the cognitive process of selectively concentrating on a discrete aspect of the communication; *emotional states* (*E*) are the feeling that we see during the communication; *intonation* (*I*) is the variations in spoken pitch that indicate the attitudes of the speaker; *self-control* (*S*) is the ability to control ones behavior and desires in communication demands; *proprioception* (*P*) represents the sense of body position

Table 1. Taxonomy of interaction types.

	Interaction		Involved technology	Cognitive processes
Explicit interaction (bidirectional)	Ocular		Camera	J, F, S
	Gesture	Facial	Camera, motion capture sensors	E, P, S
		Body	Camera, movement sensors	E, P, S
	Verbal		Microphone	E, I, U
	Tactile		Tactile sensors, accelerometer	J, S, U
	Object		Sensors, mobile devices	J, S, U
	Audio		Speakers, headphones	E, S
Implicit interaction (human to avatar)	Ocular		Camera	F
	Gesture	Facial	Camera, motion capture sensors	E, S
		Body	Camera, movement sensors	E, S
	Biofeedback interaction		Headset, vital sign devices	E, S

(e.g. hands, arms, face) related to body language; and *understanding* (*U*) is the general comprehension of matters in the conversation.

The taxonomy considers implicit and explicit interactions. Implicit interactions are related to involuntary communication elements, that is, cues such as laughing, sweating, and body postures that can reveal the inner emotional state to others.

Table 1 presents the taxonomy of interactions, including the communicative and cognitive skills that can be treated with each kind of interaction. In terms of explicit interactions using gestures, we refer to them as voluntary interactions, as they are consciously performed by the user with full intention. Because of this way to understand explicit interactions, we have classified them into the following types: Ocular, Gestural (Facial and Body-language), Verbal, Tactile, Object, and Audio Interaction.

This taxonomy, which is described in more detail in Ref. 14, aims to understand the human–avatar interactions and the way in which they affect different cognitive processes. It will also help establish reference models that guide the development of the multi-interactive avatar for SCD. Considering the conceptual prototype described in section "Design of the avatar-based interactions," we used this taxonomy as a roadmap to develop our assistive avatar that is described in the next section.

Avatar for recognizing and managing emotions

This avatar was designed with a focus on the cognitive process Emotional State (E), and it intends to assist people in the management of their emotions. This allows supporting the undertaking of cognitive function and the theory of the mind that are important research lines in SCD.[1] The avatar design is partially based on the neurological function, particularly on mirror neurons. These neurons play an important role in the acquirement of knowledge and building empathy, which is particularly useful in the case of people with SCD.

We also work with the theory of the mind as that is the capacity we have to deduce the mental state of another person. The avatar makes this information more accessible and offers an individualized training to add to the users' mental maps.

Avatar implementation

One of the implementation premises was that interacting with the avatar aids the user in better facing social communications in real life. For this reason, the avatar has a humanoid look, resembling

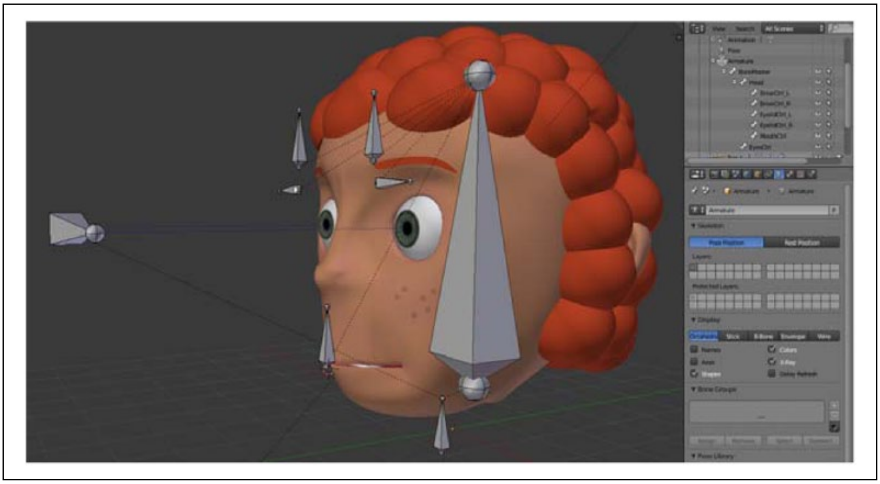

Figure 2. Final design of the avatar and the bones that control it.

reality as much as possible, and its emotional expressivity should be subtle, as it often occurs in reality. The implementation was done with Blender (https://www.blender.org/). By considering the importance of the empathy between users and the avatar gender,[8] we designed an androgynous avatar. Thus, the users themselves would decide the gender of their avatar, as to draw a more positive interaction.

After starting the design in Blender, the androgynous look has been somewhat diminished, but we have attempted to mimic it as much as possible. The design was done by extending a two-dimensional plane (in accordance with Figure 1) to a 3D model. The avatar hair was created with spheres and half-spheres to achieve the "curly" aspect. The mouth was given more attention with the design of teeth, gums, and a tongue since it has to maximize the expressivity and likeness to a human mouth. After the avatar design, we applied to it texture paint and details, finalizing with the setting of bones, which enable us to control the mesh of the avatar into moving or deforming as we wish, in this case, the expressions. Figure 2 shows the position of the bones in the avatar, as well as the end design.

As mentioned before, those bones are what enable the expressions on the avatar's face and general movement. The biggest bone controls the movement of the head (rotation and movement), the horizontal bone that is furthest from the face controls the wandering of the eyes, and the smallest bone in the bottom is the Bone Master, which is the root of the bone hierarchy and moves all the other bones if it is moved. The other bones, from top to bottom, control the eyebrows, eyelids, and mouth.

To achieve the different expressions, we modeled shape keys, which modify the position of the vertices of the object in relation to the basis. We created a shape key for each possible movement of the face, which included three different types of movement for the eyebrows (rise, frown and sad), a blinking motion, and several movements for the mouth which included the smile, frown, and two types of open mouth. This, together with the bones, allowed for the creation and movement of the expressions we deemed necessary for this model to show.

The expressions that we have included in the avatar are Happiness, Sadness, Anger, Surprise, Fear, an idle state that would consist on blinking, and a Neutral state (Figure 3). These expressions are the named universal emotion categories proposed by Paul Ekman in 1972, from which one was removed (disgust) due to it not being relevant to the conducted work.

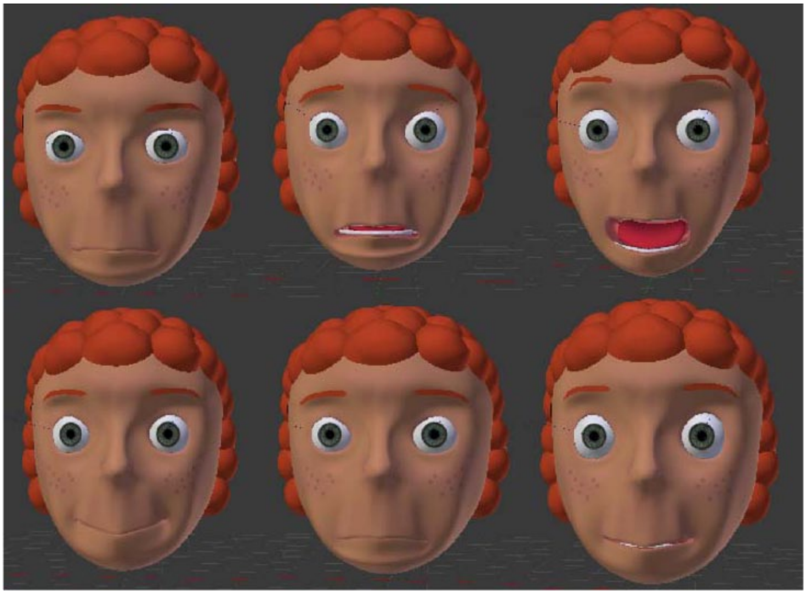

Figure 3. (Top to bottom, left to right): Anger, Fear, Surprise, Happiness, Sadness, and Neutral.

The avatar transitions between these emotions and the idle animation depending on the actions of the user, as well as the inactions. This means that if the user is in a determined state, not doing anything will also mean changing states, as it can be seen in the state machine depicted in Figure 4.

In this state machine, sliding the finger on a touch screen is considered the "Good (Action)" since in the real world this action seems to be a caress. A "Bad (Action)" is poking the touch screen since in real life it can be translated to a poke, which is more unpleasant than the caress. The term "Time (No Action)" refers to those transitions that would happen when the avatar is in certain states without receiving any user interaction, as in real life one does not stay in certain emotional states indefinitely, and this was the simplest way to mimic it. The amount of time depends on the level of arousal of the current emotion (the greater the arousal level, the smaller the timeframe to stay in that emotional state); for example, going from Surprise to Neutral would take a lot less than going from Sadness to Neutral.

Experiment to assess the emotional state of the avatar

The purpose of the experiment was to validate the 3D model of the avatar and determine whether users exhibited empathy with it. The avatar was tested by showing users the avatar going through different emotional states and asking users to report the emotional state they perceived from the avatar.

In all, 30 subjects participated in the study. They were grouped by age (C1: under 12 years; C2: 12–21 years; C3: 22–30 years, 10 people for each age group) (13 males (V) and 17 females (M)). This age range and cohorts have been determined taking into account the typical ages to diagnose and treat SCDs,[1] although none of them have any diagnosed SCDs. The experimental protocol was as follows: (1) participants were informed about the purpose of the experiment and the information to be collected; (2) the avatar in its neutral emotional state was shown to participants and the different possible interactions were explained; and (3) participants perform 20 interactions with the avatar (collecting 600 interactions during the whole experiment), and the following data were

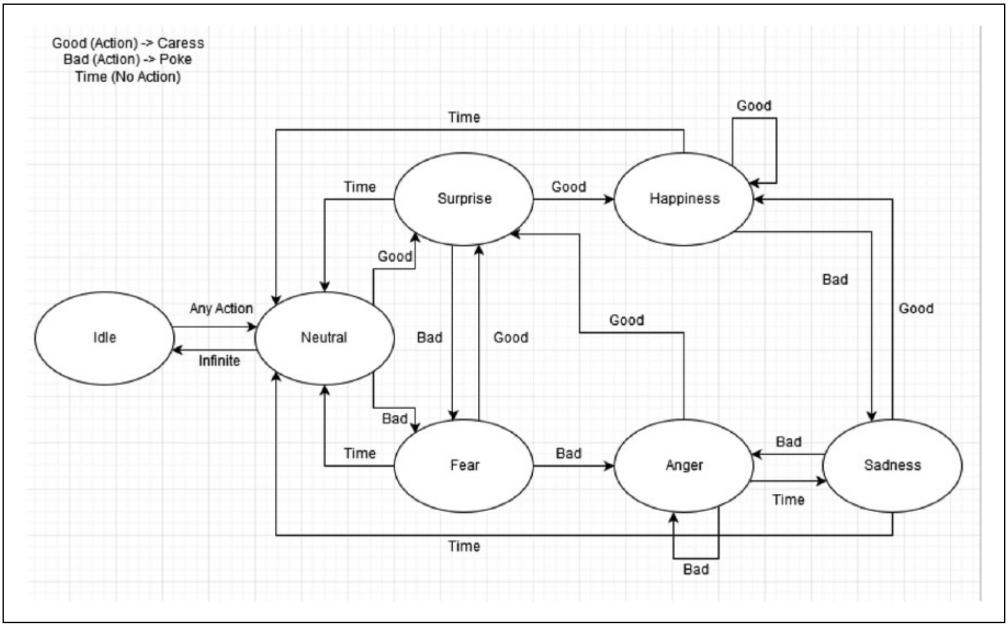

Figure 4. State machine that represents the changes in avatar emotions after certain actions.

Table 2. Accuracy in the recognition of the avatar emotions by the user (global (top) and for the 10 last interactions (bottom)).

	Neutral (%)	Fear (%)	Anger (%)	Happy (%)	Surprise (%)	Sadness (%)
Accuracy	59.88	65.26	61.90	94.16	82.05	73.68
Accuracy-10	65.15	75.56	64.10	94.27	92.00	79.31

collected: initial emotion of the avatar, final emotion of the avatar after the interaction, reflected emotion in the user (if any), and the opinion of the user about whether the emotion of the avatar is logical or not.

Regarding the avatar emotion recognition, although the users performed 20 interactions, we distinguished results on identification accuracy both using the 20 interactions and the last 10 since we used the first 10 for them to get familiar with the facial expressions.

Evaluation results

We describe the results obtained from the experiment, including accuracy, F-scores, and a confusion matrix. The mean accuracy in the correct identification of emotions during the whole test was 73.69 percent, while the mean accuracy of the correct identification of emotions only taking into account the last 10 interactions was 79.33 percent, indicating the overall improvement of correct identification overtime. Table 2 shows the recognition accuracy with all 20 interactions and the last 10. The last 10 are used since we consider the first interactions as a training period, when the user is getting familiar with the task and the avatar reactions; the Neutral emotion (59.88%; 65.15%) was the least successfully recognized mood, followed by Anger (61.90%; 64.10%) and Fear

Table 3. Global accuracy percentages by groups (C_X) and last 10 interactions (C_X-10), as well as gender.

	Neutral (%)	Fear (%)	Anger (%)	Happiness (%)	Surprise (%)	Sadness (%)
C1	79.71	78.83	59	76.67	86.67	60
C1-10	81.67	82.50	66.67	76.67	85	65
C2	51.5	62.67	35.33	95	83.83	55
C2-10	60.83	76.67	50	80	93.33	60
C3	37.5	39	81.67	97.14	65.67	87.17
C3-10	70	80	78.33	100	100	80
Male	59.30	65.38	49.49	84.62	83.72	57.70
Female	53.90	56.18	65.69	93.42	74.90	74.80

(65.26%; 75.56%). Happiness (94.16%; 94.27%) and Surprise (82.05%; 92.00%) were the most successful ones. There is a clear increase in accuracy when only the last 10 interactions are considered, indicating that the user gets familiar with the avatar.

Table 3 shows the percentages indicated in Table 2, but separated by age groups and gender.

With these numbers, we set out to calculate Recall, Precision, and F-score, which will be obtained by calculating True Positives (correct emotion identification and logical behavior), True Negatives (incorrect emotion identification and not logical behavior), False Positives (incorrect emotion identification and logical behavior), and False Negatives (correct emotion identification and not logical emotion). By logical/not logical emotion, we refer to the user's perception of the avatar's emotion being a logical/expected one or not. After gathering all this information from the tests, we have obtained the results shown in Table 4.

We can observe that the recall for Surprise of C3 is rather low, alongside the Precision of Sadness, which coincides with C2, although only the low Precision of Sadness occurs in C1, along with a low Precision for Fear. On a global level, the low Precision and Accuracy that occur in C3 and C2 are reflected in a similar manner. The confusion matrix below (Table 5) shows us the amount of mix-up when identifying an emotion, for which emotion it is confused, and whether the problem persists after the user has become more accustomed to the avatar.

We can see that most of the errors arise with the Neutral emotion, followed by confusing Fear with Sadness, Happiness with Neutral, and Sadness with Fear. Even after becoming used to the avatar, the users continue having some issues confusing Neutral and Happiness, as well as Fear and Sadness and Anger and Sadness. The likely reason for this is the similarity between Neutral and Happiness expressions, in which both look happy and welcoming, with the idea of inspiring happier emotions, but it should be rectified so that these issues do not occur again. Sadness and Anger have a similar issue, with the difference being a subtle change in the eyebrows and mouth, for which we will strive to make them clearly different. The mix for Fear confused for Sadness is very possibly the frowns on both faces, and we will seek to rectify this as well. Finally, the interactions themselves and the feelings that have appeared the most are shown in Tables 6 and 7.

As it can be seen, most users chose to caress the avatar, followed by poking and then waiting. Younger users (group C1) engaged more in general, forgoing the wait most of the time and choosing to caress. This was because after observing the avatar behavior, they figured out that caressing made the avatar happy, and they seemed to have preferred seeing the avatar in this emotional state, even choosing not to do anything other than this action in the later part of the experiment. Groups C2 and C3, aged 12–30 years, have a much more even distribution of the actions, with an overall small difference among the three groups.

Table 4. Global, C1, C2, and C3 values for TP, TN, FP, and FN, as well as the calculation of Recall, Precision, and F-score for every emotion.

	Neutral	Fear	Anger	Happiness	Surprise	Sadness
Global						
TP	70	53	34	121	59	37
TN	14	9	1	5	4	15
FP	8	13	10	27	7	39
FN	15	11	4	7	39	6
Recall	0.82	0.83	0.89	0.95	0.60	0.86
Precision	0.90	0.80	0.77	0.82	0.89	0.49
F-score	0.86	0.82	0.83	0.88	0.72	0.62
Global-C1						
TP	7	8	16	76	24	22
TN	1	1	0	0	1	2
FP	1	5	6	5	0	18
FN	0	1	1	1	1	1
Recall	1	0.89	0.94	0.99	0.96	0.96
Precision	0.88	0.62	0.73	0.94	1	0.55
F-score	0.93	0.73	0.82	0.96	0.98	0.70
Global-C2						
TP	26	20	10	20	22	7
TN	5	5	0	3	2	4
FP	4	4	1	10	6	15
FN	6	5	0	4	16	4
Recall	0.81	0.80	1	0.83	0.58	0.64
Precision	0.87	0.83	0.90	0.67	0.79	0.32
F-score	0.84	0.82	0.95	0.74	0.67	0.42
Global-C3						
TP	38	22	12	27	9	8
TN	6	3	1	2	0	8
FP	2	2	3	11	1	7
FN	12	5	3	2	22	1
Recall	0.76	0.81	0.80	0.93	0.29	0.89
Precision	0.95	0.92	0.80	0.71	0.90	0.53
F-score	0.84	0.86	0.80	0.81	0.44	0.67

TP: True Positives; TN: True Negatives; FP: False Positives; FN: False Negatives.

The most visited state was "Happiness" and "Neutral," although the latter is due to the fact that it is the initial emotion. For all groups, happiness was the emotion more frequently elicited, while Sadness is the least visited emotion in the two youngest groups, whereas Anger is the least visited emotion for the older group, followed closely by Fear.

Not counting Neutral, the users had overall a more positive than negative stream of emotions (56.05% were positive, 43.95% were negative, although the split for the youngest age group was more: 59.14%–41.86%). This is a positive result, and the likelihood of the small difference is due to the older age groups commenting on the fact that feelings such as "Surprised" on the avatar annoyed them when they wanted to make the avatar happy.

Overall, most users had a positive response to the avatar, and they commented that the feeling of "Surprise" puzzled them after a certain amount of interactions, citing that "it shouldn't feel

Table 5. Global confusion matrix (right) and for the last 10 interactions (left).

Actual/predicted-10	N	Fe	An	Ha	Su	Sa	Actual/predicted	N	Fe	An	Ha	Su	Sa
Neutral	43	0	0	15	3	5	Neutral	98	13	1	28	10	12
Fear	0	34	1	0	1	9	Fear	0	62	1	1	3	26
Anger	4	1	25	0	0	11	Anger	7	1	39	0	0	16
Happiness	4	0	0	67	0	0	Happiness	11	0	0	129	0	0
Surprise	0	1	1	0	46	0	Surprise	1	12	2	5	97	0
Sadness	0	1	4	0	1	23	Sadness	4	1	7	0	2	42

N: Neutral; Fe: Fear; An: Anger; Ha: Happiness; Su: Surprise; Sa: Sadness.

Table 6. Total actions done by users and by groups (Poke (unpleasant), Caress (pleasant), and Wait (no action)).

	Poke	Caress	Wait
C3	64	70	66
C2	72	70	58
C1	69	113	18
Total	205	253	142

Table 7. Total number of emotions the avatar has shown, divided by groups.

	Neutral	Fear	Anger	Happiness	Surprise	Sadness
C1	73	34	20	34	36	13
C2	62	38	23	24	48	16
C3	27	23	20	79	33	28
Emotions	162	95	63	137	117	57

surprised after a while." Another reason for the small difference is that we are not accounting for Neutral, which was a fairly common feeling for older age groups. Upon observing the given answers, most children were fairly polarizing, and if the avatar was happy or surprised, they were as well, whereas negative emotions made them feel Sadness more than Fear or Anger. Sadness also had a high incidence of erroneous identifications; therefore, we will strive to fix these issues on the model.

Conclusion and future work

This work contributes to the field of health-related affective systems, particularly to help design avatar-based interactive applications to evaluate and assist people with SCD. First, the taxonomy presented in section "Taxonomy of human–avatar interactions" can aid the research community when developing interactive systems using avatars, providing a better understanding of both the interactions and the way in which they affect particular cognitive processes. Second, being the main contribution of this article, we have evaluated interactions with affective avatars, and we have learnt the following lessons:

- Since the design objective of the avatar was to show emotions in a very subtle way, as they occur in real life, the evaluation highlighted the need to redesign some of them, due to low recognition accuracy by the normotypical users.
- We identified a significant improvement in emotion identification after the first 10 interactions; thus, it is important to give the user certain time to get used to the avatar and distinguish its emotions and behavior.
- We obtained evidence that the current state machine of affective behavior makes sense for the users when they successfully identify the avatar's emotions.
- Results evidence that users have a mimic response of emotions; they generally react with positive emotions (mainly happiness) to positive states of the avatar and also feel sad when the avatar exhibits a negative emotion.
- Some differences among age cohorts have been identified. For example, we observed that C1 (less than 12-year-old users) forgoes wait without doing any interaction and performs more pleasant interactions than unpleasant.

In terms of discussion, the learnt lessons seem to evidence characteristic interaction patterns that show empathic behavior of users, for example, to avoid repetitive negative interaction in order to avoid unpleasant avatar responses and a mimic response to emotions. These findings are more evident in C2 and C3 cohorts. These patterns can enhance future assistive avatars in terms of SCD diagnosis and treatment, mainly with children over 12 years old. Currently, and after redesigning some of the avatar emotions, we are collecting more evaluation data with normotypical users and people with SCD related to empathy and emotion management (mainly Down's syndrome and autism spectrum disorders). These data can establish a more valuable ground truth about empathic interactions with avatars and would allow the training of learning algorithms to identify empathic and non-empathic behaviors.

In terms of future work, in order to maximize the *human–avatar interaction*, we are including an aspect of emotion recognition so as to gather a more precise reaction from the avatar to the facial expressions of the user. This will be accomplished using active shape models and support vector machines that have proven to yield good results.[17]

Acknowledgements

This work was conducted in the context of UBIHEALTH project under International Research Staff Exchange Schema (MC-IRSES 316337).

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This work was partially supported by FONDECYT (Chile), grant no. 1150252.

References

1. Gibson J, Adams C, Lockton E, et al. Social communication disorder outside autism? A diagnostic classification approach to delineating pragmatic language impairment, high functioning autism and specific language impairment. *J Child Psychol Psychiatry* 2013; 54(11): 1186–1197.
2. LoPresti EF, Mihailidis A and Kirsch N. Assistive technology for cognitive rehabilitation: state of the art. *Neuropsychol Rehabil* 2004; 14(1–2): 5–39.

3. Hanke S, Sandner E, Stainer-Hochgatterer A, et al. The technical specification and architecture of a virtual support partner. In: *AmI'15, European conference on ambient intelligence*, Athens, 11–13 November 2015.
4. Adjouadi M, Sesin A, Ayala M, et al. Remote eye gaze tracking system as a computer interface for persons with severe motor disability. In: Miesenberger K, Klaus J, Zagler W, et al. (eds) *Computers helping people with special needs* (LNCS, vol. 3118). Berlin: Springer, 2004, pp. 761–769.
5. Plischke H and Kohls N. Keep it simple! Assisting older people with mental and physical training. In: Stephanidis C (ed.) *Universal access in human-computer interaction: addressing diversity* (LNCS, vol. 5614). Berlin: Springer, 2009, pp. 278–287.
6. Ugarte A, García I, Ortiz A, et al. User interfaces based on 3D avatars for interactive television. In: Cesar P, Chorianopoulos K and Jensen JF (eds) *Interactive TV: a shared experience* (LNCS, vol. 4471). Berlin: Springer, 2007, pp. 107–115.
7. Norris AE, Weger H, Bullinger C, et al. Quantifying engagement: measuring player involvement in human–avatar interactions. *Comput Hum Behav* 2014; 34: 1–11.
8. Inal Y, Sancar H and Cagiltay K. Children's avatar preferences and their personalities. In: *Society for information technology & teacher education international conference*, Orlando, FL, USA, 20–25 March 2006.
9. Cereghetti D, Kleanthous S, Christophorou C, et al. Virtual partners for seniors: analysis of the user's preferences and expectations on personality and appearance. In: *AmI'15, European conference on ambient intelligence*, Athens, 11–13 November 2015.
10. Malhotra A, Hoey J, König A, et al. A study of elderly people's emotional understanding of prompts given by Virtual Humans. In: *10th international conference on pervasive computing technologies for healthcare*, Cancun, Mexico, 16–19 May 2016.
11. Iglesias A and Luengo F. AI framework for decision modeling in behavioral animation of virtual avatars. In: Shi Y, van Albada GD, Dongarra J, et al. (eds) *Computational science* (LNCS, vol. 4488). Berlin: Springer, 2007, pp. 89–96.
12. Bailenson JN, Yee N, Merget D, et al. The effect of behavioral realism and form realism of real-time avatar faces on verbal disclosure, nonverbal disclosure, emotion recognition, and copresence in dyadic interaction. *Presence Teleop Virt* 2006; 15(4): 359–372, http://dx.doi.org/10.1162/pres.15.4.359
13. Kang S and Watt J. The impact of avatar realism and anonymity on effective communication via mobile devices. *Comput Hum Behav* 2013; 29(3): 1169–1181.
14. Johnson E, Hervás R, Mondéjar T, et al. Improving Social Communication Disorders Through Human-Avatar Interaction. In: *Ambient Intelligence for Health*, Puerto Varas, Chile, 1–4 December 2015, pp. 237–243.
15. McCallum S. Gamification and serious games for personalized health. *St Heal T* 2012; 177: 85–96.
16. King D, Greaves F, Exeter C, et al. Gamification: influencing health behaviours with games. *J Roy Soc Med* 2013; 106(3): 76–78.
17. Lozano-Monasor E, López MT, Fernández-Caballero A, et al. Facial expression recognition from webcam based on active shape models and support vector machines. In: Pecchia L, Chen L, Nugent C, et al. (eds) *Ambient assisted living and daily activities* (LNCS, vol. 8868). Berlin: Springer, 2014, pp. 147–154.

Article

Health Informatics Journal

Health Informatics Journal
2018, Vol. 24(2) 194–205

DOI: 10.1177/1460458216663023
journals.sagepub.com/home/jhi
SAGE

Screening diabetes mellitus 2 based on electronic health records using temporal features

Angela Pimentel
Universidade Nova de Lisboa, Portugal

André V Carreiro
INESC-ID, Portugal

Rogério T Ribeiro
APDP ERC, Portugal

Hugo Gamboa
Universidade Nova de Lisboa, Portugal

Abstract
The prevalence of type 2 diabetes mellitus is increasing worldwide. Current methods of treating diabetes remain inadequate, and therefore, prevention with screening methods is the most appropriate process to reduce the burden of diabetes and its complications. We propose a new prognostic approach for type 2 diabetes mellitus based on electronic health records without using the current invasive techniques that are related to the disease (e.g. glucose level or glycated hemoglobin (HbA1c)). Our methodology is based on machine learning frameworks with data enrichment using temporal features. As as result our predictive model achieved an area under the receiver operating characteristics curve with a random forest classifier of 84.22 percent when including data information from 2009 to 2011 to predict diabetic patients in 2012, 83.19 percent when including temporal features, and 83.72 percent after applying temporal features and feature selection. We conclude that he pathology prediction is possible and efficient using the patient's progression information over the years and without using the invasive techniques that are currently used for type 2 diabetes mellitus classification.

Keywords
classification, database, diabetes mellitus 2, electronic health record, prognostic tool, screening

Corresponding author:
Angela Pimentel, LIBPhys-UNL, Departamento de Física, Faculdade de Ciências e Tecnologia, Universidade Nova de Lisboa, 2829-516 Caparica, Portugal.
Email: a.pimentel@campus.fct.unl.pt

Introduction

About 347 million people worldwide have diabetes. In 2012, diabetes was the direct cause of 1.5 million deaths.[1] The prevalence of diabetes is increasing in developed and developing countries and is predicted to achieve 7.7 percent worldwide by 2030.[2]

Type 2 diabetes mellitus (T2DM) is the most common form of diabetes. In T2DM, the central problem seems to be insulin resistance in peripheral tissues. This resistance is favored by genetic susceptibility and is made worse by weight gain and lack of exercise. The pancreas compensates for the lack of insulin action on cell glucose uptake by producing more insulin. Eventually, not even this hyperinsulinemia is enough to avoid the rise in blood glucose levels. To make matters worse, pancreas exhaustion leads later to loss of secretory capacity and to full-blown diabetes symptoms (e.g. excess thirst, frequent urination, and constant hunger). T2DM emerges and progresses slowly, remaining undetected for years.[3]

Chronic hyperglycemia of diabetes is associated with long-term dysfunction, damage, and failure of various organs, especially the eyes, kidneys, nerves, heart, and blood vessels. Individuals with undiagnosed T2DM are also at significantly higher risk of stroke, coronary heart disease, and peripheral vascular disease than the non-diabetic population.[4]

Treatment of T2DM prevents some of its devastating complications[5,6] but does not usually restore normoglycemia or eliminate all the adverse consequences. The diagnosis is often delayed until complications are present, and therefore, early screening for diabetes or prediabetes assumes a great importance.[7]

Screening for blood glucose has been used as a tool to identify individuals at high risk of developing diabetes or already with asymptomatic diabetes. There is debate regarding whether screening for fasting glucose is sufficient or an oral glucose tolerance test is needed for detection of asymptomatic diabetes.[8] Measuring either fasting or postprandial blood glucose is an invasive procedure, as well as costly and time-consuming. Blood glucose has a large random variation and only gives information on a subject's current glycemic status. Furthermore, the true primary prevention would be to identify high-risk subjects when they are still in a normoglycemic state and to treat them by interventions that prevent their transition from normoglycemia to impaired glucose tolerance and to overt diabetes.[9]

Electronic health records (EHRs) are a systematic collection of electronic health information about an individual patient or population. It is a mechanism for integrating healthcare information on an electronic medical record with the purpose of improving quality of care.[10] One of its main potentials is enhancing decision support systems and facilitating reliable and reproducible outcomes of research and reporting.[11,12] With machine learning techniques, the challenge is to extract meaningful information from data, to gain knowledge and patterns of discovery.[13]

This article proposes a new prognostic approach for T2DM given an EHR and without using the current invasive techniques, such as glucose level and glycated hemoglobin (HbA_{1c}). This approach, based on machine learning methods, supports the applicability of using the patient's EHR to efficiently and non-invasively predict patients who are likely to develop T2DM. The result is a screening tool that allows the identification of high-risk subjects in the population, increases awareness of risk factors the pathology presents, and also emphasizes the importance of using EHR. The entire dataset captures 4 years (2009–2012) of "live" visits, and a T2DM prediction of diabetes diagnosis incidence by 2012 can be performed, with 84.22 percent of area under the receiver operating characteristics curve (AUC) using a random forest (RF) classifier. It has been shown that T2DM can be efficiently prevented by lifestyle modifications.[14–16] Hence, our tool aims to help to identify individuals who would benefit from this intensive lifestyle counseling.

Material and methods

Diabetes dataset

The diabetes database is available on Kaggle,[17] a community of data scientists who compete to solve complex data science problems. This EHR dataset was provided by the Practice Fusion community, who released training and test sets of de-identified, Health Insurance Portability and Accountability Act (HIPAA)-compliant medical records to spur innovation into new uses of clinical data to improve public health and patient care. This dataset is one of the largest and richest sources of medical record data ever released and includes information on diagnoses, lab results, medications, allergies, immunizations, vital signs, and health behavior.

Practice Fusion sponsored the following challenge with this dataset: "Identify patients diagnosed with T2DM." This article proposes a different approach: "Predict patients who will present T2DM in the following time period T+1, based on previous data from T." Our goal is to predict who will develop T2DM instead of identifying those who already have been diagnosed with T2DM, resulting in a prognostic prediction instead of diagnosis. A number of details from the dataset were removed, and cannot be accessed, in order to avoid the manipulation of diabetes prediction since they are direct measures related to the disease. Specifically, the diagnosis feature of T2DM patients as defined by *International Classification of Diseases, 9th Revision* (ICD-9) codes, diabetes medication, glucose levels, insulin, and HbA_{1c} were removed. For this competition, an indicator column was added designating who has a diagnosis of T2DM. This diagnosis is assigned to each patient over the 4 years of acquired data. Since T2DM is a chronic disease, this label can be used to predict patients in 2012 given the information available from previous years. The disadvantage, since we cannot access the T2DM patient's diagnosis and year of diagnosis, is that we are limited to provide a screening analysis over 1 year. However, we can still provide a temporal analysis of the available parameters for this time window.

Feature extraction. We proceeded to a feature extraction based on the state-of-the-art descriptions available on the Kaggle[17] diabetes community. The following steps were performed: (1) outliers were cleaned and height median calculated for each patient; (2) body mass index (BMI) was recalculated with this constant height for each patient, eliminating the noise of measured fluctuations; (3) for the features weight, height, BMI, systolic blood pressure, diastolic blood pressure, temperature, and respiratory rate, the median was calculated and truncated to maximum and minimum; (4) the diagnostic features were grouped at two levels based on clinical etiology or symptoms similarities according to the ICD-9 codes: 245 Level 2 and 22 Level 1 groups were created; (5) the medication features were treated at the active principle level. For each National Drug Code (NDC) active principle, administration route and dose were extracted. After that, active principles were grouped into families taking into account chemical similarities or common clinical indication. Features were created for maximum dose of an active principle/family, number of active principles in the family administered to the patient, number of prescriptions, and binary flags; (6) features such as "Smoking Status" and "Previous Smoking Status" were created; (7) other features like allergy or immunization were not considered due to low number of patients.

In the end, the data comprise information from 2009 to 2011 with the label representing the status in year 2012, to allow a screening analysis. The training set will be represented by 70 percent of the original data and the testing set with the remaining data (30%). There are 529 features for 9947 patients (43% are male and 57% are female) with ages between 21 and 93 years. Overall, 19 percent of the patients are diabetic, and stratification was ensured when splitting the training and testing sets, guaranteeing similar class distributions in both sets. The programming languages used for this processing were R and Python. Training was performed using the Amazon Elastic Compute Cloud (Amazon EC2) with an m3.2xlarge model (CPU 8 cores and 30 GB memory).[18]

Classification

After preprocessing the data, we proceeded to the diabetes screening by assessing different classifiers. We used two machine learning packages for this analysis: Weka[19] and Theano.[20] With the Weka package, the standard classifiers were used, initially with the default parameters. These include Naïve Bayes (NB),[21] alternating decision tree (ADT),[22] RF,[23] random tree (RT), k-nearest neighbor (KNN),[24] and support vector machine (SVM) with a polynomial kernel implementation.[25] The predictive models were built using 5×10-fold cross-validation (CV) on the training data.[26] The classifier's parameters were optimized using a meta-classifier that performs CV given a list of possible parameters.[27] Also, knowing that the classes are unbalanced (19% are diabetic), we used another technique available on Weka—Synthetic Minority Over-sampling Technique (SMOTE). This technique is applied in each CV train fold and will be responsible for over-sampling the minority class by creating synthetic examples, therefore allowing the analysis of a more balanced training data.[28] We chose to use a SMOTE of 0, 150, and 300 percent, where 300 percent means that for each minority class instance, the algorithm creates another three synthetic ones, resulting in a train set with an approximated 50-50 distribution.

Using Theano, it is possible to attain speeds rivaling hand-crafted C implementations for problems involving large amounts of data. The classifier Multilayer Perceptron was used with Theano due to its low learning speed on big data, and a different number of hidden neurons were tested (250, 500, and 1000).

Temporal features. It has been shown that T2DM is associated with weight gain[29,30] and raised blood pressure.[31] Therefore, the weight, BMI, and systolic and diastolic blood pressure evolution analysis over time may improve our predictive model. We applied linear regression of the chosen features over the years (2009–2011), and we used the slope and interception as the new temporal features (TF), a data enrichment which is used to investigate possible improvements on classification. We have selected this type of regression for two reasons: (1) clinician feedback that supports that temporal trajectories would be probabilistically linear for a general population; (2) it is highlighted by Adler et al.,[32] Hanson et al.,[33] and Bays et al.[34] that the temporal trajectories of the selected features are linear. However, as future work, we intend to complement this study by analyzing other types of regressions (e.g. quadratic).

Feature selection. Given the high number of used features in this study (529), we performed a feature selection (FS) analysis in order to verify whether some of the original features would be discarded as being irrelevant to the classification. The InfoGainAttribute evaluator was tested among a variety of FS algorithms available on Weka and provided better AUC results. This algorithm evaluates the worth of an attribute by measuring the information gain with respect to the class attribute. Concisely, the information gain is a measure of the reduction in entropy of the class variable after the value for the feature is observed. The Ranker search method was used retaining the top 150 attributes ranked by their individual evaluations. Nevertheless, we consider that as future

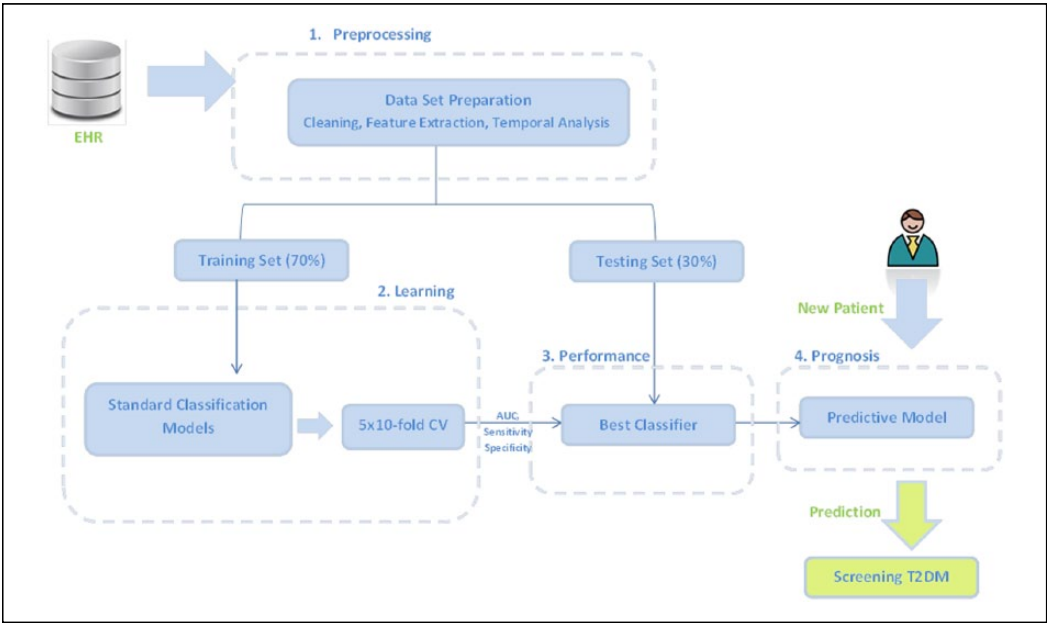

Figure 1. Summary of the followed workflow. Based on EHR, the data are prepared and divided in training and testing sets. Different classifiers are applied on the training set with a stratified 5 × 10-fold CV and the best model is chosen based on its metrics (AUC, sensitivity, and specificity). Later, its performance is evaluated using the testing set, and a final predictive model is built allowing a prognosis of T2DM given a new patient's information.

work a feature engineering analysis should be considered in order to improve our classification system.

Figure 1 summarizes the workflow used for the analysis. We tested different classifiers on the original data (OD), on the original data with the inclusion of temporal features (OD + TF), and later with feature selection (OD + TF + FS). Regarding the performance evaluation, several metrics were retrieved, such as the receiver operating characteristic (ROC) area (AUC), sensitivity, and specificity. Afterwards, the best classifier was chosen based on the best metrics obtained and applied on the test data. Our predictive model will then allow a prognostic prediction of T2DM when a new patient's information is given. The programming languages used for this analysis were Java for the Weka packages and Python for Theano.

Results and discussion

In this section, we present and discuss the results of the predictive models on the training set (70% of initial dataset). Afterwards, we show the performance evaluation of the best predictive model on the test set (remaining 30%).

Training set

The results are shown in Tables 1 and 2, for the AUC, specificity and sensitivity, respectively, after applying 5 × 10-fold CV on the training set for different classifiers.

Table 1. AUC values obtained from the train set in percentage (stratified 5 × 10-fold CV).

AUC (%)							
SMOTE	Data	RF	NB	SVM	ADT	RT	KNN
0%	OD	**85.38 ± 0.11**	**81.56 ± 0.05**	65.59 ± 0.23	**81.81 ± 0.3**	61.11 ± 1.1	64.48 ± 0.37
	OD + TF	**85.13 ± 0.1**	**81.54 ± 0.05**	65.93 ± 0.22	**81.8 ± 0.28**	62.66 ± 0.64	63.01 ± 0.73
	OD + TF + FS	**85.52 ± 0.13**	**81.72 ± 0.04**	65.76 ± 0.26	**81.8 ± 0.28**	63.17 ± 0.51	63.61 ± 0.38
150%	OD	**86.45 ± 0.05**	72.78 ± 0.51	73.66 ± 0.17	**81.14 ± 0.16**	62.29 ± 0.53	62.64 ± 0.61
	OD + TF	**85.32 ± 0.12**	71.77 ± 0.3	66.38 ± 0.48	**80.81 ± 0.57**	60.95 ± 0.85	59.62 ± 0.27
	OD + TF + FS	**85.52 ± 0.16**	79.45 ± 0.39	67.71 ± 0.11	**81.37 ± 0.44**	63.59 ± 0.6	61.83 ± 0.31
300%	OD	**86.2 ± 0.07**	70.95 ± 0.2	75.27 ± 0.29	**80.55 ± 0.15**	62.37 ± 0.34	63.68 ± 0.66
	OD + TF	**85.13 ± 0.16**	68.72 ± 0.46	66.94 ± 0.53	79.71 ± 0.57	61.17 ± 0.8	59.7 ± 0.34
	OD + TF + FS	**85.24 ± 0.12**	79.03 ± 0.29	69.69 ± 0.15	**80.5 ± 0.24**	63.05 ± 0.31	62.1 ± 0.26

AUC: area under the receiver operating characteristics curve; CV: cross-validation; SMOTE: Synthetic Minority Over-sampling Technique; RF: random forest; NB: Naïve Bayes; SVM: support vector machine; ADT: alternating decision tree; RT: random tree; KNN: k-nearest neighbor; OD: original data; TF: temporal features; FS: feature selection; Bold values represent AUC values higher than 80 percent.

Table 2. Specificity and sensitivity values obtained from the train set in percentage (stratified 5 × 10-fold CV).

SMOTE	Data	RF	NB	SVM	ADT	RT	KNN
Specificity (%)							
0%	OD	99.25 ± 0.05	78.74 ± 0.12	97.32 ± 0.11	95.95 ± 0.27	87.83 ± 0.56	92.64 ± 0.18
	OD + TF	99.24 ± 0.08	78.69 ± 0.12	97.13 ± 0.14	95.95 ± 0.27	88.09 ± 0.53	92.83 ± 0.26
	OD + TF + FS	97.36 ± 0.12	79.11 ± 0.07	98.49 ± 0.12	96.26 ± 0.24	88.3 ± 0.32	92.58 ± 0.4
150%	OD	97.06 ± 0.04	76.82 ± 0.17	87.98 ± 0.1	87.42 ± 0.94	82.67 ± 0.11	85.37 ± 0.27
	OD + TF	99.04 ± 0.08	77.38 ± 0.72	95.48 ± 0.27	93.59 ± 0.43	85.27 ± 0.46	88.88 ± 0.26
	OD + TF + FS	96.81 ± 0.1	83.69 ± 0.11	93.05 ± 0.32	91.32 ± 0.61	84.82 ± 0.25	85.01 ± 0.34
300%	OD	95.3 ± 0.1	58.09 ± 0.39	80.55 ± 0.29	80.62 ± 0.92	81.05 ± 0.3	82.9 ± 0.17
	OD + TF	98.92 ± 0.11	60.56 ± 1.9	94.48 ± 0.39	93.61 ± 0.64	85.23 ± 0.24	88.05 ± 0.43
	OD + TF + FS	96.1 ± 0.2	84.04 ± 0.24	89.67 ± 0.49	88.3 ± 0.48	84.2 ± 0.42	82.91 ± 0.09
Sensitivity (%)							
0%	OD	14.16 ± 0.26	65.98 ± 0.26	33.86 ± 0.51	30.01 ± 0.88	31.95 ± 1.91	25.34 ± 0.49
	OD + TF	12.81 ± 0.33	66.42 ± 0.21	34.73 ± 0.51	30.01 ± 0.88	32.33 ± 1.02	23.75 ± 0.78
	OD + TF + FS	20.3 ± 0.57	66.67 ± 0.09	34.28 ± 0.55	30.01 ± 0.88	35.15 ± 0.86	24.31 ± 0.28
150%	OD	32.52 ± 0.41	62.39 ± 1.1	59.34 ± 0.33	49.55 ± 1.65	41.77 ± 0.96	39.88 ± 0.8
	OD + TF	16.29 ± 0.39	60.89 ± 1.08	37.28 ± 1.14	34.41 ± 2.22	35.07 ± 1.63	29.97 ± 0.97
	OD + TF + FS	29.42 ± 0.28	56.17 ± 0.38	42.37 ± 0.43	41.47 ± 2.09	40.22 ± 1.16	38.8 ± 0.32
300%	OD	40.42 ± 0.48	74.08 ± 0.57	70.0 ± 0.38	61.58 ± 1.31	43.69 ± 0.86	44.27 ± 0.81
	OD + TF	16.28 ± 0.7	67.89 ± 1.02	39.39 ± 1.39	31.56 ± 2.57	34.93 ± 1.39	31.02 ± 0.77
	OD + TF + FS	31.53 ± 0.45	55.98 ± 0.73	49.7 ± 0.68	46.38 ± 1.07	39.85 ± 0.77	41.41 ± 0.56

CV: cross-validation; SMOTE: Synthetic Minority Over-sampling Technique; NB: Naïve Bayes; SVM: support vector machine; RF: random forest; ADT: alternating decision tree; RT: random tree; KNN: k-nearest neighbor; OD: original data; TF: temporal features; FS: feature selection.

Using Theano, the accuracy is the only metric available for the Multilayer Perceptron (MP) classifier (with 500 hidden neurons), and 5 × 10-fold CV was not applied due to package specifications. The results for the accuracy were in the order of 81.08 ± 0.15 percent, and given that this was outperformed by other classifiers, the results for this classifier were discarded, suggesting as future work a deeper understanding of neural networks using this approach.

Table 3. Feature ranking of the top 20 selected features.

Feature selection ranking—Top 20	
1. Hypertension essential	11. Active principle
2. Year of birth	12. Min body mass index
3. High/low blood pressure	13. Number of prescripts
4. Mixed hyperlipidemia	14. Total diagnostics per weighted year
5. Statin (no. of prescriptions)	15. Number of 3 digit groups of diagnostics with medication
6. Statin (dose adjusted)	16. angiotensin-converting-enzym inhibitor(ACEI) (no. of prescriptions)
7. Statin (no. of different active principles)	17. ACEI (bin)
8. Max body mass index	18. Median systolic blood pressure
9. Max systolic blood pressure	19. Total diagnostics
10. Median body mass index	20. Max weight

The first feature, "Hypertension Essential," is considered the most important for classification, and the "Max Weight" is considered the less important in this list.

The RF classifier provides good results for AUC and specificity values. However, the sensitivity outcomes are not as good as the previous metrics. This means that given a patient that is likely to be diabetic, the classifier may fail to assert this prediction. The best sensitivity outcomes are returned by the NB classifier. However, our best classifier will be chosen based on the best combination of all the three metrics. The low sensitivity values seen in Table 2 may also be justified by the unbalanced classes that represent the training set (19% are diabetic), which explains the use of the SMOTE filter in this analysis.

Our main research goal was to show that diabetes screening could be made accurately and that this screening would be efficient using the patient's information over the years. We notice that, for the RF classifier, the AUC values are improved with the use of a SMOTE filter of 150 percent, although better results could be expected for a SMOTE of 300 percent where the classes are balanced. Later, the addition of TF and FS did not improve the AUC values in comparison with the OD. The best specificity values are found when no synthetic instances are created (SMOTE 0%), and addition of TF slightly improves the specificity values. Still for the RF classifier, the sensitivity results are improved with the increasing SMOTE percentage. And the OD individually outperforms the use of TF and FS on the sensitivity values.

Using FS, we notice that sensitivity values are improved with the RF classifier and that two temporal features (weight slope and interception) are included in these ranked features. We also see that the most important features for the classification are essential hypertension, year of birth, systolic and diastolic blood pressure, and mixed hyperlipidemia (Table 3). These features are in agreement with the known risk factors of the disease.[35–37] Although the Kaggle competition uses a different proposal analysis, our most important features are also in agreement with the Kaggle competitors.

These ranked factors can also be used to alert the physician about the main parameters that should be considered when screening T2DM. The diagnostics on stress and rheumatic pain are found to be the less important features for the classification.

Following the RF, ADT also provides good results for AUC values. Given that decision trees implicitly perform feature selection and are easy to interpret, we consider that Figure 2 provides important insights for screening T2DM for this population. A new instance is scored by summing all of the prediction nodes through which it passes.

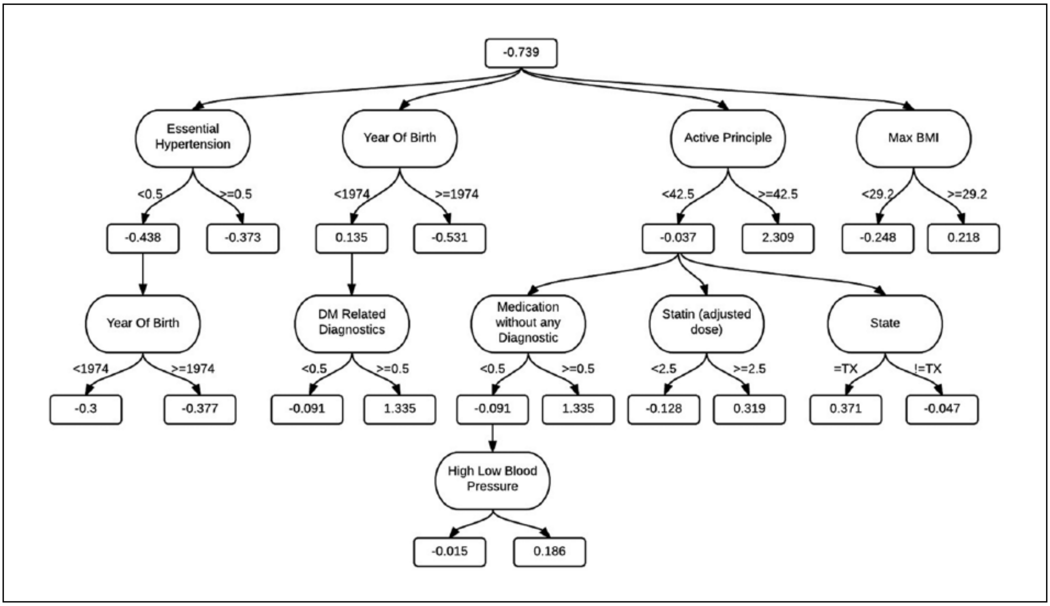

Figure 2. ADT visualization. A prognosis using this model only requires nine features. A new instance is scored by summing all of the prediction nodes through which it passes. If this final sum is positive, then it is classified as a T2DM prognostic.
Source: Adapted from the Weka Explorer Software.
DM: diabetes mellitus; TX: Texas state.

Statistical comparisons were also made between each group with a Wilcoxon signed-rank test as suggested from Demsar.[38] The results show that there are no statistical significant differences when adding TF or applying FS for a SMOTE of 0 percent. However, for a SMOTE of 150 and 300 percent, there are statistical significant differences after using TF, decreasing the classifiers performance, especially for the SVM classifier. We also found that there are no statistical significant differences between each SMOTE group. This suggests that more advanced techniques for unbalanced datasets should be considered for future work as well as a deeper understanding of the temporal trajectories and feature selection analysis.

In order to compare the performance of the different classifiers, we applied the Friedman test, suggested by Demsar,[38] using the IBM SPSS Statistics Editor. The results show that there are statistically significant differences between the classifiers for the AUC values. We analyzed the pairs comparison, with significance values corrected for multiple testing, and found that the RF and ADT performed significantly better than RT, KNN, and SVM ($p \leqslant 0.03$). Also, NB significantly outperformed KNN and RT. Moreover, we verified that the top performing classifier, according to the mean test rank, was RF. Given that there were no statistical significant differences between each SMOTE group, and knowing that the best AUC outcomes with the RF classifier occur with a SMOTE filter of 150 percent, we conclude that the best option is to proceed with this filter. The average training time processing for these datasets using the RF classifier with a SMOTE of 150 percent, with parameter optimization, and using stratified 5×10-fold CV was approximately 2 h (with Amazon EC2).

We have chosen the RF due to its performing metrics outputs and statistical analysis. However, this classifier has other advantages: it runs efficiently on large datasets, minimizes overfitting, is

Table 4. Results in percentage obtained from the test set using the RF classifier.

RF				
SMOTE	Data	AUC	Sensitivity	Specificity
0%	OD	83.08	11.64	98.99
	OD + TF	82.88	11.48	99.12
	OD + TF + FS	83.19	16.07	99.28
150%	OD	**84.22**	29.19	96.42
	OD + TF	**83.19**	14.1	98.7
	OD + TF + FS	**83.72**	26.52	96.21
300%	OD	84.11	36.23	93.77
	OD + TF	83.03	13.45	98.91
	OD + TF + FS	83.39	29.84	95.92

RF: random forest; SMOTE: Synthetic Minority Over-sampling Technique; AUC: area under the receiver operating characteristics curve; OD: original data; TF: temporal features; FS: feature selection; Bold values represent AUC values higher than 80 percent.

fast, scalable, and does not require tuning many parameters (such as SVM). For these reasons, we proceeded to the test set with the RF classifier.

Test set

After choosing the best classifier based on the training set, we analyzed its performance on the test set. The entire training set was used on the RF to build our predictive model, and then the test set was used and the AUC, sensitivity, and specificity metrics were retrieved as seen in Table 4. These results will allow us to know what to expect with the RF prediction, when dealing with new and unknown patient's information.

Comparing the results with the training set, we find that the test metrics are below the confidence interval (mean ± standard deviation), with exception to the specificity values for the OF + TF with a SMOTE of 300 percent and OD + TF + FS with a SMOTE of 0 percent. However, it is a slight deviation that should not compromise our analysis.

Another interesting fact for the AUC outcomes is that the best performing values occur with a SMOTE filter of 150 percent, which is in agreement with our training conclusions.

Although the use of TF and FS did not improve the final performance, we still believe that these models are important and useful methods in prognostic analyses. Finally, the simple theory, fast speed, stability, and insensitivity to noise of the RF classifier make it suitable for screening T2DM using EHR.

Although FS did not select gender as one of the most important features, we also tested these classification models separately for male and female. We found out that the best performance metrics using a RF classifier with a SMOTE of 150 percent for male are below the expected (AUC = 78%), although female values are similar when compared using both genders (AUC = 84%). This is a topic that we wish to address as future work.

Conclusion

In this article, we propose a new approach of T2DM prognostic prediction, given a set of previous EHR. During the preprocessing, we have added TF based on the state-of-the-art knowledge of the disease risk factors. FS was also applied given the high number of available features. We applied

5×10-fold CV on the training set for different classifiers and selected the classifier that presented the best metrics: AUC, sensitivity, and specificity. Following this, the test set was applied on the best model—the RF classifier. We achieved an AUC performance of 84.22 percent using the data comprising the years 2009–2011 to predict diabetic patients in 2012, and 83.19 and 83.72 percent after the inclusion of TF and TF with FS, respectively.

The inclusion of both TF and FS did not improve model performance. This strengthens the importance of a prospective study for patient temporal progression trajectories and feature selection analysis. Nonetheless, we strongly believe that these approaches should be considered in a pathology prognosis—even more for T2DM, a chronic disease that will get progressively worse if left untreated. The major drawback of this study was the impossibility of accessing the T2DM diagnosis date. This limited us to only provide a prognosis of 1 year. Also, it is of our main interest to understand which patients do not have the disease from 2009 to 2011 and were diagnosed with T2DM on 2012. These patients are the ones who should have an increased attention so that we can understand which features are more affected with disease onset. Nevertheless, our predictive model presents promising results for screening T2DM using EHR, which could be used when the development of the disease is still at an early stage. As for future work, we intend to apply this methodology on different datasets; we wish to study the temporal trajectories for this population and also test other feature selection algorithms. We also wish to address the T2DM prognosis based on gender, which may enhance our understanding of T2DM progression. Finally, we wish to classify new patients with our predictive model and monitor their evolution over a year to confirm our preliminary screening.

Declaration of conflicting interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) received no financial support for the research, authorship, and/or publication of this article.

References

1. WHO—World Health Organization. January 2015, http://www.who.int/en/
2. Shaw JE, Sicree RA and Zimmet PZ. Global estimates of the prevalence of diabetes for 2010 and 2030. *Diabetes Res Clin Pract* 2010; 87: 4–14.
3. Brill MT. *Diabetes*. Minneapolis, MN: Twenty-First Century Books, 2011.
4. Association AD. Screening for diabetes. *Diabetes Care* 2002; 25: 21–24.
5. UK Prospective Diabetes Study Group (UKPDS). Intensive blood-glucose control with sulphonylureas or insulin compared with conventional treatment and risk of complications in patients with type 2 diabetes. *Lancet* 1999; 354: 602.
6. UK Prospective Diabetes Study Group (UKPDS). Effect of intensive blood-glucose control with metformin on complications with overweight patients with type 2 diabetes. *Lancet* 1998; 352: 1558.
7. Harris MI and Eastman RC. Early detection of undiagnosed diabetes mellitus: a US perspective. *Diabetes Metab Res Rev* 2000; 16: 230–236.
8. Will new diagnostic criteria for diabetes mellitus change phenotype of patients with diabetes? Reanalysis of European epidemiological data. *Brit Med J* 1998; 317: 371–375.
9. Lindström J and Tuomilehto J. The diabetes risk score a practical tool to predict type 2 diabetes risk. *Diabetes Care* 2003; 26(3): 725–731.
10. Gunter TD and Terry NP. The emergence of national electronic health record architectures in the United States and Australia: models, costs, and questions. *J Med Internet Res* 2005; 7: e3.

11. Australian council for safety and quality in health care, open disclosure standard: a national standard for open communication in public and private hospitals, following an adverse event in health care 2003. *Former Australian Council for Safety and Quality in Health Care*, July 2003.
12. Donsa K, Spat S, Beck P, et al. Towards personalization of diabetes therapy using computerized decision support and machine learning: some open problems and challenges. In: Holzinger A, Roecker C and Ziefle M (eds) *Smart health, lecture notes in computer science LNCS 8700*. Heidelberg; Berlin: Springer, 2015, pp. 235–260.
13. Holzinger A, Dehmer M and Jurisica I. Knowledge discovery and interactive data mining in bioinformatics—State-of-the-art, future challenges and research directions. *BMC Bioinformatics* 2014; 15(S6): I1.
14. Pan X, Li G, Hu Y, et al. Effects of diet and exercise in preventing NIDDM in people with impaired glucose tolerance: the Da Qing IGT and diabetes study. *Diabetes Care* 1997; 20(4): 537–544.
15. Tuomilehto J, Lindstrom J and Eriksson JG. Prevention of type 2 diabetes mellitus by changes in lifestyle among subjects with impaired glucose tolerance. *N Eng J Med* 2001; 344(18): 1343–1350.
16. Knowler WC, Barrett-Connor E, Fowler SE, et al. Reduction in the incidence of type 2 diabetes with lifestyle intervention or metformin. *N Eng J Med* 2002; 346(6): 393–403.
17. Kaggle: Go from big data to big analytics. October 2014, https://www.kaggle.com/
18. AWS | Amazon elastic compute cloud. (EC2)—scalable cloud hosting, October 2014, http://aws.amazon.com/
19. Hall M, Frank E and Holmes G. The WEKA data mining software: an update. *ACM SIGKDD Explor Newsl* 2009; 11(1): 10–18.
20. Bergstra J, Breuleux O, Bastien F. Theano: a CPU and GPU math expression compiler. In: *Proceedings of the Python for Scientific Computing Conference (SciPy)*, 2010, http://www.iro.umontreal.ca/~lisa/pointeurs/theano_scipy2010.pdf
21. John GH and Langley P. *Estimating continuous distributions in Bayesian classifiers* (UAI'95). San Francisco, CA: Morgan Kaufmann Publishers, 1995, pp. 338–345.
22. Freund Y and Mason L. The alternating decision tree learning algorithm. *Int Conf Mach Learn* 1999; 99: 124–133.
23. Breiman L. Random forests. *Mach Learn* 2001; 45(1): 5–32.
24. Aha DW, Kibler D and Albert M.K. Instance-based learning algorithms. *Mach Learn* 1991; 6(1): 37–66.
25. Platt J. Sequential minimal optimization: a fast algorithm for training support vector machines. Technical report msr-tr-98-14, 1 April 1998.
26. Kohavi R. A study of cross-validation and bootstrap for accuracy estimation and model selection. In: *Proceedings of the 14th international joint conference on Artificial intelligence* (vol. 14), 20 August 1995, pp. 1137–1145. San Francisco, CA: Morgan Kaufmann Publishers, http://web.cs.iastate.edu/~jtian/cs573/Papers/Kohavi-IJCAI-95.pdf
27. Kohavi R. *Wrappers for performance enhancements and oblivious decision graphs* (Technical report). Stanford, CA: Stanford University, 1995.
28. Chawla NV, Bowyer KW, Hall LO, et al. SMOTE: synthetic minority over-sampling technique. *J Artif Intell Res* 2002; 16: 321–357.
29. Colditz GA, Willett WC, Rotnitzky A, et al. Weight gain as a risk factor for clinical diabetes mellitus in women. *Ann Intern Med* 1995; 122(7): 481–486.
30. Chan JM, Rimm EB, Colditz GA, et al. Obesity, fat distribution, and weight gain as risk factors for clinical diabetes in men. *Diabetes Care* 1994; 17(9): 961–969.
31. Kannel WB and McGee DL. Diabetes and cardiovascular disease: the Framingham study. *JAMA* 1979; 241(19): 2035–2038.
32. Adler AI, Stratton IM, Neil HAW, et al. Association of systolic blood pressure with macrovascular and microvascular complications of type 2 diabetes (UKPDS 36): prospective observational study. *BMJ* 2000; 321(7258): 412–419.
33. Hanson RL, Narayan KMV, McCance DR, et al. Rate of weight gain: weight fluctuation, and incidence of NIDDM. *Diabetes* 1995; 44(3): 261–266.

34. Bays HE, Chapman RH and Grandy S. The relationship of body mass index to diabetes mellitus, hypertension and dyslipidaemia: comparison of data from two national surveys. *Int J Clin Pract* 2007; 61(5): 737–747.
35. Watkins P. *ABC of diabetes*. 5th ed. London: BMJ Books, 2003.
36. Fonseca V, Pendergrass M and Harrison McDuffie R. *Diabetes in clinical practice.* Springer, 2010.
37. Bilous R and Donnelly R. *Handbook of diabetes.* John Wiley & Sons, 2010.
38. Demsar J. Statistical comparisons of classifiers over multiple datasets. *J Mach Learn Res*. 2006; 7: 1–30.

Health Informatics Journal

Article

Health Informatics Journal
2018, Vol. 24(2) 206–215

DOI: 10.1177/1460458216663024
journals.sagepub.com/home/jhi
SAGE

Vital sign documentation in electronic records: The development of workarounds

Jean E Stevenson
The University of Sheffield, UK; Linnaeus University, Sweden

Johan Israelsson
Kalmar County Hospital, Sweden

Gunilla Nilsson and Goran Petersson
Linnaeus University, Sweden

Peter A Bath
The University of Sheffield, UK

Abstract
Workarounds are commonplace in healthcare settings. An increase in the use of electronic health records has led to an escalation of workarounds as healthcare professionals cope with systems which are inadequate for their needs. Closely related to this, the documentation of vital signs in electronic health records has been problematic. The accuracy and completeness of vital sign documentation has a direct impact on the recognition of deterioration in a patient's condition. We examined workflow processes to identify workarounds related to vital signs in a 372-bed hospital in Sweden. In three clinical areas, a qualitative study was performed with data collected during observations and interviews and analysed through thematic content analysis. We identified paper workarounds in the form of handwritten notes and a total of eight pre-printed paper observation charts. Our results suggested that nurses created workarounds to allow a smooth workflow and ensure patients safety.

Keywords
electronic health records, healthcare professionals, patient safety, vital signs, workarounds

Corresponding author:
Jean E Stevenson, eHealth Institute, Linnaeus University, Kalmar, Sweden SE 391 82.
Email: jean.stevenson.agren@lnu.se

Introduction

Workarounds can be defined as non-standard methods for accomplishing work blocked by dysfunctional processes[1] and are common in complex environments such as healthcare settings. Dealing with unexpected situations is common practice for healthcare professionals, and thus, they are often masters at workarounds.[2] The introduction of electronic health records (EHRs) in healthcare settings may have led to an escalation of workarounds.[2]

Increasingly used for all aspects of clinical documentation, EHRs have the potential to improve the quality of health care by facilitating communication and the accessibility of information.[3] Problems with EHRs have also been identified: lack of overview, excessive time to document and retrieve information and poor navigability.[4] Because of these problems, EHRs do not meet the needs of front-line staff who, in turn, often resort to workarounds in order to achieve their goals.[5] One problematic area is in the documentation of vital signs: for example, temperature, pulse, respiratory rate, blood pressure and oxygen saturation.[6] This is a matter of concern as close 'monitoring' of vital signs is key to early recognition of clinical deterioration in hospital patients. Studies have repeatedly shown that early detection of clinical deterioration and timely management can reduce the incidence of adverse events: for example, cardiac arrest, unplanned admission to intensive care units (ICUs), prolonged hospital stay or death.[7–12] A previous study demonstrated a lack of completeness of vital sign documentation in an EHR in patients who had suffered cardiac arrest.[6]

Current study

This study builds on our previous work by looking more closely at work processes and the flow of information regarding the measurement, documentation and retrieval of vital signs and investigates the extent to which the EHR supports documentation of these. A combination of observational and interview methods was used to collect the data. The aim of this article is to identify and describe workarounds that were related to vital signs.

Methods

This study was conducted using a qualitative approach.[13] Data were collected between May and September 2014.

Setting

The research setting was a district general hospital in Sweden with 372 beds. Three clinical settings were involved: the cardiology department consisting of two wards, a cardiac ICU and a high dependency unit; an infection ward; and an emergency department. The hospital began using an EHR in 2007.

Data collection

There were two separate, but linked, methods used for data collection: an observational study and an interview study. An observational study can be a rich source of information as it enables the researcher to capture what people do rather than what they say they do[14] and promote understanding of complex situations.[15] Semi-structured interviews were conducted to augment the data collected during observations and provide an in-depth understanding of views and experiences of the personnel.[13]

Observational study

In the observational study, one of the researchers (J.E.S.) spent 62 h shadowing around 15 nurses and recording detailed information about all aspects of vital sign measurement, documentation and retrieval. All three shifts were observed to increase the reliability of the data by ensuring that any variations in practice from one shift to another were recorded. Handwritten notes were taken in the field. As a supplement to observations, opportunistic interviews were conducted. This meant that the researcher could ask follow-up questions during observations in less busy moments to clarify what was being observed and gain direct views.[16]

Semi-structured interviews

Following the observational study, 13 semi-structured interviews were conducted to investigate the views and experiences of doctors (n = 3) and nurses (n = 10). All interviews were audio-recorded, and supplementary handwritten notes were taken. The audio-recordings were translated and transcribed (J.E.S.) and a second researcher (G.N.) validated the accuracy of the translation and transcription.

Data analysis

To analyse the observational data, content analysis was carried out whereby themes emerged from the coded data. These themes were used to structure interview questions.[17] Content analysis can reduce large volumes of material into content categories.[18] Throughout the analysis, field notes from observations and interview transcripts were used to corroborate one another.

Ethical issues

Ethical approval was granted from the Central Ethical Review Board, Linköping, Sweden. Permission to perform the observational study was sought from the managers of the departments involved. Information about the study was given to potential participants. The nurses and doctors who volunteered gave their consent to participate in the study and were informed that they could withdraw at any time. Anonymity of the participants was guaranteed.

Results

A total of 62 h of observations were carried out in three separate clinical fields. In addition, 13 medical and nursing staff were interviewed. The age range of those interviewed was 27–55 years. The mean length of interview was 19 min, and work experience ranged from <1 to 20 years.

Since workarounds are deviations from expected work practices, the accepted standard practice for documentation and retrieval of vital signs is described: to document vital signs in the EHR in a table called ‘measurements’, from which these could also be retrieved.

Workarounds

We identified several paper workarounds related to the documentation of vital signs. The workarounds were in the form of notes written on scraps of paper, ‘post-it’ notes and pocket notebooks. Paper observation charts had also been created to suit the clinical requirements of each of the three clinical areas. In the cardiology unit, there were three paper charts; in the infection ward, there were four paper charts; and in the emergency department, one paper chart. The results from each of these three areas are presented separately, in order to illustrate the diversity of the workarounds adopted.

Cardiology

One of the workarounds in cardiology was handwritten notes on pieces of paper or in notebooks. Frequently, the EHR was not beside the patients when vital signs were recorded. The field note and interview excerpts below illustrate this:

> The nurse measures the vital signs, writes the measurements in a notebook, then goes to the trolley in the corridor and transfers her notes on vital signs from the notebook to the measurement table in the EHR. (Field note 2, Cardiac ward, Night shift 06:00)

> We often double document. Most often we write on a paper because it is easier to have a paper with you beside the patient, than the big trolley with the computer on it. You have a little paper where you write down all the vital signs. Then you go out and write it in in the table for measurements in the EHR. (Interview, Nurse 5)

The examples above suggest that the proximity of the EHR in relation to the patient played a role in the use of paper notes. Usually, vital signs from notes were entered into the EHR directly after the nurse left the room, but if there was an interruption, or something more important came up, entry into the EHR could be delayed until the end of the shift indicating possible time lapses and the potential for mistakes to be made. The fact that it was 'easier' to write on paper may imply usability issues with the EHR. The problem of double documentation, that is, the same information being recorded in separate places, was recognised by the staff as a risk because of potential error during transcription.

The following interview excerpt describes how nurses would like to work to avoid such problems:

> Instead of writing it down on a paper, if you could maybe put them on an iPad and then you could send it over to the EHR, as they are so little and simple … It would save time for me and also the more often you write the same thing the higher the risk is that it can be wrong. (Interview, Nurse 2)

Another workaround was the use of paper charts: there were three paper observation charts used to support the documentation of vital signs. These were in the format of a table. First, there was an observation chart for patients who were recovering from a coronary angiography or a percutaneous coronary intervention (PCI). This chart was used after these procedures to check several factors including the insertion site, distal pulse and blood pressure. The following field note is an example of its use:

> Patient returns from angiography. BP and pulse checked frequently – 5, 10 and 15 minutes then half hourly. The nurse explains that with these frequent recordings, the table in the EHR would become far too long. Therefore a paper chart is used and scanned into the EHR when the patient is discharged. (Field note 4, HDU)

A second observation chart was specifically for observing patients who had been prescribed certain medications, such as a nitroglycerine infusion. These interview comments illustrate its purpose:

> Well if I started up a nitroglycerine infusion, maybe I'd take 10 blood pressures, but I would never put 10 blood pressures in the EHR, I would enter a few of them to show how it was when we started … (Interview, Nurse 3)

The third observation chart was for all other patients who needed frequent vital signs, for example, if their condition had deteriorated or had become unstable. Here, pulse, blood pressure, respiratory rate, oxygen saturation and urinary output could be recorded. The following interview excerpts illustrate its use:

> Then it can be that you have an acutely sick patient and you don't have time to document all this (in the EHR). Then it is paper charts that are suitable. (Interview, Nurse 5)

> On paper it is certainly easier to see. Then you don't need to click in the EHR, … you can see it quickly. (Interview, Nurse 6)

> Then we enter the first and the final BP in the EHR instead of filling in each pulse and blood pressure … (Interview, Nurse 1)

The examples above have several implications: first, that the table in the computer could not accommodate frequent vital sign recordings; second, that it took less time to document on paper; third, that it took less time to view vital signs on paper; and finally, that it was easier to write by hand on paper than to use the EHR.

During interviews, nurses described documenting vital signs as time-consuming and cumbersome, for example, it took approximately 20 'clicks' to enter one blood pressure in the EHR. For doctors, retrieval of vital signs was problematic as the following interview excerpts describe:

> It is very difficult to get to this information from the EHR … I must make many clicks to see those vital signs … They are not easily accessible. (Interview, Cardiology consultant)

A second doctor said that it was difficult to get an overview of the patient's vital signs and that it took many clicks:

> If a patient deteriorates, it's quicker to ask the registered nurse (RN) what the vital signs are than it is to find them in the EHR. (Interview, Cardiology doctor)

Nurses often remembered the latest vital signs or could glance at their paper notes to answer questions about vital signs from the doctors. This suggests that there may be usability problems with the EHR, and the necessity to adopt another workaround – verbal communication.

Infection ward

There were four paper charts in the infection ward, two in the form of tables and two as 'fill-in' forms. The first of these was similar to the observation chart in the cardiology ward for frequent vital signs. A registered nurse (RN) explained this:

> I mean, you don't write everything in the EHR when it might be 10/12 times a day. Then we have a paper that is later scanned in. We would write some recordings in the EHR, e.g., when we finish a shift … there is a paper for this, just so you can write it in quickly. Maybe you are measuring the signs every 15 minutes

> … you maybe add the highest sign and the lowest sign to the EHR … We don't have time to go to the EHR every time … There would be an awful lot if we wrote all the recordings … in the table (meaning the table in the EHR). (Interview, Nurse 8)

In this instance, the paper chart allowed vital signs to be entered quickly. There could be a time lapse before vital signs appeared in the EHR. Furthermore, all of the vital signs recorded were not documented in the EHR. This suggests that the table in the EHR was not suitable for frequent documentation. The need for this paper chart was confirmed by an infection ward doctor at interview:

> If we have a sepsis patient who is unstable in the ward, then there is a paper chart which is left in the room with the patient, so that you don't need to run back and forth between the computer and the patient. You have it close by and then you can use that chart instead … maybe taking vitals every 15 minutes during several hours. It can then be valuable to have it like this … especially in an infection ward when you maybe have an infectious patient. (Interview, Doctor 3)

This shows that it was important for the documentation of vital signs to be close to the patient to allow both easy entry and easy retrieval of information.

The second paper chart in the infection ward was the 'temp list'. The routine was to record temperatures twice daily, at 06:00 and 18:00. At 06:00, temperatures were written on the 'temp list' and added to the EHR afterwards. The following field note indicates what happened next:

> At 07.30 the 'temp list' is on the trolley, along with a laptop which is being used for administering medications. The nurse checks the oxygen saturation of a patient which the night nurse had said was labile. She adds the oxygen saturation to the paper 'temp list'. After she has been round all the patients, she sits down at the EHR and transfers the new information from the paper 'temp list' to the EHR. (Field note 9, Nurse, Day shift)

Thus, the paper 'temp list' was used to enable the most previous record of vital signs to be at the nurse's fingertips. Although called a 'temp list', additional vital signs were sometimes noted here if necessary. The following interview excerpts describe why a 'temp list' was considered necessary:

> (The RN shows me a temperature list). With us, all patients have their temperature checked. That's what we are interested in and then we BAS the patients too [this means to check blood pressure, respiratory rate and oxygen saturation], then write it on the temp list too. (The RN shows me how more vital signs have been added in small handwriting at the side of the temperature slot). (Interview, Nurse 8)

> We have the list so that we have it when we go round (the patients). You don't always have time to stand at the computer. We write it in later. It's easier to write it on a paper and then put it in the EHR when you have time. We have a routine that it is written in the EHR before we go home. (Interview, Nurse 8)

This indicates that nurses found it easier to write on paper and that it saved time. Doctors found it necessary to ask nurses about vital signs as noted in this quote:

> First I ask the nurse for the latest vital signs. They can often be given orally. (Interview, Doctor, Infection ward)

Verbal communication was, therefore, another workaround to overcome the problem of delayed entry of vital signs to the EHR. An RN's statement suggested reasons for workarounds:

> … it (documentation) has to fit in with how we work … go smoothly. We can't have it that we stand at the computer the whole time and only write things. It's more important that we look after the patients, then document when we have time. Sure, we must document, but *when* we document is not the most important, just that everything is documented. (Interview, Nurse 8, italics added for emphasis)

This suggests that official documentation in the EHR was considered something that just had to be done, while paper workarounds were included in the actual process of care, and that the priorities of the staff for recording important information were different from those for which the EHR was developed.

There were two further paper forms which had been created specially by the infection ward personnel. These were used during admission, and each included vital signs.

Emergency department

The emergency department had only one paper form. It was a form used for writing down details of patients when they were admitted and included vital signs. The following field note illustrates its use:

> A paper form is used to write the vital signs, with boxes for temperature, pulse, blood pressure, respiratory rate, oxygen saturation. Then the signs are added to the EHR. The paper form was put in a plastic folder and placed at the nurses' station. A nurse related that it was not really needed but that the doctors liked to have these vital signs on a piece of paper, and with them when they examined a patient. (Field note 8)

This suggests that doctors found it more convenient to have a paper form at hand rather than to retrieve vital signs from the EHR; it allowed a quick overview of the patient which the doctor could take a paper with him or her to the patient's bedside.

Discussion

In this investigation, we found several workarounds that were directly linked to the flow of information relating to vital signs. There were eight paper-based workarounds. Some verbal workarounds also emerged. Previous studies have recognised that paper use can persist after implementation of electronic systems if they do not support clinical workflow.[1,19] Essentially, clinical staff had created these workarounds because they felt that it would ensure patient safety and the flow of normal work processes. The justifications and consequences of workarounds are discussed in this section based on four aspects: technical, operational, cultural and organisational, in line with Ser et al.[20]

Technical aspects

EHR designers' lack of understanding of work processes, for example, in monitoring a patient's vital signs, led to the development of an inadequate system. Consequently, lack of adequate technological design led to workarounds in the form of paper charts for documenting vital signs because of the perceived efficiency of paper over the corresponding function in the computer.[16] In the EHR, there was a bespoke table for vital signs, which could accommodate routine, infrequent documentation. However, when more frequent vital signs were required, the table was considered unsuitable. Staff complained of excessive clicking to enter vital signs, so it was quicker to write by hand.

Moreover, the paper chart had space for as many vital signs as it was necessary to record, and it was easy to view at a glance implying, as noted in previous studies,[4,16] that there were problems with usability. To sum up, the reasons for workarounds were problems with accessibility, visibility and readability of the data in the EHR. In addition, efficient and timely entry of data was not possible as electronic documentation could not be carried out at the point of care because of the cumbersome equipment.

These findings may explain results from a previous study in which the vital signs were incomplete in the EHR.[6] When staff used paper workarounds, only some of the vital signs documented on paper were added to the EHR. This may have caused fragmentation and inconsistency in the electronic record.[19] In addition, nurses frequently reported that vital signs were only added to the EHR 'later' or at the end of a shift. Delayed entry, as well as the use of paper charts, had implications for staff who wanted to check vital signs in the EHR. Doctors, in particular, expressed frustration that vital signs were so difficult to find and to compound this problem, when they located vital signs they could not be sure these were the latest. Thus, a further workaround emerged in the form of verbal communication – doctors asking nurses for the latest vital signs.

Operational factors

A laptop was used where possible to record information in the EHR. However, it was not always possible to take a laptop into the room, for instance, if no laptop was available, if a patient was being barrier nursed or if there was other equipment needed at the bedside, such as a blood pressure machine. Therefore, nurses used handwritten paper notes for vital signs or charts, which were kept at the bedside if patients required more frequent recordings. Later, some of these vital signs were entered into the EHR, which introduced the issue of double documentation and risk for error during transcription, as well as delayed entry in some cases. Thus, from the staff perspective, the EHR did not integrate with their existing work practices.

Cultural factors

The workarounds used in two of the three clinical areas revealed a kind of subculture in relation to the types of patients cared for. For instance, in the infection ward, a paper chart used twice daily was called a 'temperature list', presumably because of the relationship between infection and fever. The chart had only one slot for recording vital signs, that is, for temperature. If other vital signs were thought necessary, these were squeezed in at the side of the temperature slot. As this happened quite often, it is perhaps surprising that a more suitable chart with space for supplementary vital signs had not been designed. In the cardiology unit, observation charts also reflected that specific vital signs were considered more important than others, for example, blood pressure was the focus in the post-angiography chart. Although these examples show initiative and resourcefulness on the part of clinical staff, as they strive to ensure safe patient care despite an inappropriate system, these home-spun paper charts were not in accordance with current thinking on patient monitoring, which recommends the recording of complete sets of vital signs in order to detect clinical deterioration.[21] In addition, research on paper observation charts emphasises the importance of plotted graphs and colour coding to maximise the cognitive ability of users,[22] but these features were not apparent in the home-spun, tailor-made paper charts used on these wards. Paper charts may be an appropriate workaround,[19] but it could be important that high-quality paper charts be used if maximum benefit is to be gained.

Organisational issues

It has been shown that EHR systems often serve the needs of strategic and managerial users but do not meet the needs of front-line workers.[5] Nurses have alerted managers to problems of documentation in the EHR but perceived that they were not listened to[4] or that inadequate solutions were initiated to alleviate the problems they encountered. For example, the 'measurements table' was a vast improvement on the tool used in a previous study[6] but still did not meet the requirements of vital sign documentation. Therefore, it is not surprising that workarounds persist and that there are homemade solutions at ward level. When systems are not adequate, workers address problems head on, and over time, solutions become part of the work routine.[23] However, organisations have a responsibility to guarantee patient safety and should not depend on front-line workers to solve problems created by inadequate technological design. The downside of workarounds is that although they work well for front-line staff, and become embedded as routines, they may cause organisational inertia, and the reason they were needed in the first place becomes forgotten. This can mean that organisations do not learn from or solve these inadequacies which, in the longer term, can be inefficient and expensive.[23]

Limitations of the study

This study provided valuable insights, although it is important to note that there are limitations to be considered in relation to the findings. The study took place in one setting and examined one EHR system so that the results may not be transferable to all settings. However, the study identified workarounds that staff found necessary to circumvent problems of an EHR widely used in Sweden and could provide the basis for similar studies in other locations.

Conclusion

Our study suggests that nurses created workarounds to get the job completed efficiently and provide their patients with safe care responsibly. There was an air of acceptance that the EHR was the way that it was, their voices had not been heard and they had found their own solutions. To be safe, electronic record systems should be correct, consistent and current,[19] but the risk of error during transcription or incomplete and delayed entries make it increasingly unlikely that these positive benefits will be realised. On one hand, resilient front-line staff may find ways to serve their information needs and mitigate potential perceived risks to patient safety by creating workarounds in the form of tailor-made paper charts. On the other hand, locally developed paper charts may not necessarily be based on current evidence and instead be based on tacit knowledge.[24] Conclusively, organisations and system providers should embrace system design flaws as learning opportunities to improve patient safety and efficiency.

Acknowledgements

The authors would like to acknowledge the cooperation of the study hospital for allowing this research to be carried out. They also want to thank the staff who participated in the observations and interviews. Furthermore, the authors would like to thank the Information School, University of Sheffield, UK, and the eHealth Institute, Linnaeus University, Kalmar, Sweden, for supporting this study.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) received no financial support for the research, authorship and/or publication of this article.

References

1. Tucker AL. Workarounds and resiliency on the front lines of health care. In: Wachter RMEA (ed.) *Morbidity and mortality rounds in the web*. Cambridge, MA: Harvard University, 2009.
2. Zhou X, Ackerman MS and Zheng K. CPOE workarounds, boundary objects, and assemblages. In: *CHI 2011 - 29th annual CHI conference on human factors in computing systems, conference proceedings and extended abstracts*, 2011, pp. 3353–3362. DOI: 10.1145/1978942.1979439.
3. Ayatollahi H, Bath PA and Goodacre S. Factors influencing the use of IT in the emergency department: a qualitative study. *Health Informatics J* 2010; 16: 189–200.
4. Stevenson JE and Nilsson GC. Nurses' perceptions of an electronic patient record from a patient safety perspective: a qualitative study. *J Adv Nurs* 2012; 68: 667–676.
5. Eason K and Waterson P. Fitness for purpose when there are many different purposes: who are electronic patient records for? *Health Informatics J* 2014; 20: 189–198.
6. Stevenson JE, Israelsson J, Nilsson G, et al. Recording signs of deterioration in acute patients: the documentation of vital signs within electronic health records in patients who suffered in-hospital cardiac arrest. *Health Informatics J*, 22(1): 21–33.
7. Bristow PJ, Hillman KM, Chey T, et al. Rates of in-hospital arrests, deaths and intensive care admissions: the effect of a medical emergency team. *Med J Aust* 2000; 173: 236–240.
8. Bellomo R, Goldsmith D, Uchino S, et al. A prospective before-and-after trial of a medical emergency team. *Med J Aust* 2003; 179: 283–287.
9. Jones D, Bellomo R, Bates S, et al. Long term effect of a medical emergency team on cardiac arrests in a teaching hospital. *Crit Care* 2005; 9: 808–815.
10. Buist MD, Moore GE, Bernard SA, et al. Effects of a medical emergency team on reduction of incidence of and mortality from unexpected cardiac arrests in hospital: preliminary study. *BMJ* 2002; 324: 387–390.
11. McQuillan P, Pilkington S, Allan A, et al. Confidential inquiry into quality of care before admission to intensive care. *BMJ* 1998; 316: 1853–1858.
12. Hillman KM, Bristow PJ, Chey T, et al. Antecedents to hospital deaths. *Intern Med J* 2001; 31: 343–348.
13. Punch KF. *Introduction to social research: quantitative and qualitative approaches*. London: SAGE, 2005.
14. Wisker G. *The postgraduate research handbook*. Basingstoke: Palgrave, 2001.
15. Bowling A. *Research methods in health*. 3rd ed. Maidenhead: McGraw Hill, 2009.
16. Saleem J, Russ A, Neddo A, et al. Paper persistence, workarounds, and communication breakdowns in computerized consultation management. *Int J Med Inform* 2011; 80: 466–479.
17. Elo S and Kyngäs H. The qualitative content analysis process. *J Adv Nurs* 2008; 62: 107–115.
18. Polit DF and Beck CT. *Nursing research: generating and assessing evidence for nursing practice*. 8th ed. Philadelphia, PA: Lippincott Williams & Wilkins, 2008, p. 796.
19. Russ A, Saleem J, Justice C, et al. Electronic health information in use: characteristics that support employee workflow and patient care. *Health Informatics J* 2010; 16: 287–305.
20. Ser G, Robertson A and Sheikh A. A qualitative exploration of workarounds related to the implementation of national electronic health records in early adopter mental health hospitals. *PLoS ONE* 2014; 9: e77669.
21. DeVita MA, Smith GB, Adam SK, et al. Identifying the hospitalised patient in crisis: a consensus conference on the afferent limb of rapid response systems. *Resuscitation* 2010; 81: 375–382.
22. Horswill M, Preece M, Hill A, et al. *Detecting abnormal vital signs on six observation charts: an experimental comparison*. Sydney, NSW, Australia: Australian Commission on Safety and Quality in Health Care, 2010, pp. 1–25.
23. Tucker AL. The impact of operational failures on hospital nurses and their patients. *J Oper Manag* 2004; 22: 151–169.
24. Kobayashi M, Fussell SR, Xiao Y. Work coordination, workflow, and workarounds in a medical context. In: *CHI'05 Extended abstracts on human factors in computing systems*, 2 April 2005, pp. 1561–1564. ACM.

Article

Health Informatics Journal

Health Informatics Journal
2018, Vol. 24(2) 216–224

DOI: 10.1177/1460458216667166
journals.sagepub.com/home/jhi
SAGE

Barriers for recruitment of patients with chronic obstructive pulmonary disease to a controlled telemedicine trial

Eva Broendum, Charlotte Suppli Ulrik, Thorbjorn Gregersen, Ejvind Frausing Hansen, Allan Green and Thomas Ringbaek
Hvidovre Hospital, Denmark

Abstract

The aim of this analysis is to investigate reasons why patients with chronic obstructive pulmonary disease decline to participate in a controlled trial of telemedicine. Patients with previous chronic obstructive pulmonary disease exacerbations were invited to participate in a 6-month randomized telemedicine trial. For eligible patients, reasons for refusal were registered. Of 560 eligible patients, 279 (50%) declined to participate in the trial, 257 (92%) reported a reason: 53 (20.6%) technical concerns, 164 (63.8%) personal reasons, 17 (6.6%) preferred outpatient clinic visits, and 23 (8.9%) did not want to participate in clinical research. Compared to consenting patients, subjects declining participation were significantly older, more often female, had higher lung function (%predicted), lower body mass index, higher admission-rate for chronic obstructive pulmonary disease in the previous year, and were more often diagnosed with osteoporosis. Many eligible patients decline participating in a controlled tele-healthcare trial and, furthermore, a tailored approach for recruiting females and elderly patients appears appropriate.

Keywords

assistive technologies, decision-support systems, ehealth, organizational change and information technology, telecare

Introduction

Telemedicine (TM) has in recent years gathered a lot of attention as a way to improve monitoring of patients' condition. Particularly within the area of chronic obstructive pulmonary disease (COPD), home monitoring has been seen as a means to prevent hospital admissions and thus to improve the efficiency in treatment and time spent on transportation.[1] The research in this area, however, has not yet reached conclusive results; some studies have shown reductions in admission[2–5] whereas others have found no effect.[6–9] The effects of TM are insufficiently clarified,

Corresponding author:
Eva Broendum, Hvidovre Hospital, Kettegaard Alle 30, 2650 Hvidovre, Denmark.
Email: eva.broendum@regionh.dk

and currently there are more than 30 randomized studies reported to www.clinicalTrials.gov. Despite mixed evidence of effectiveness and cost-effectiveness, a political decision has been made to implement TM to patients with COPD in Denmark. Thus, recruitment issues are important to explore.

Patient recruitment to TM trials has also been reported as difficult. However, more detailed studies into the reasons for refusal of participation, bias in selection, and drop-out are lacking.[10] A recent review found a mean refusal rate of 33 per cent (4%–71%) in patients with heart failure or COPD, the most common reason for refusal being 'not interested and/or believing TM to be unnecessary'.[10] Studies into the rates and reasons for refusal among patients with COPD – as the primary outcome of interest – have not yet been reported. However, based on the available studies the rate of non-participation seems to vary from 22 to 80 per cent.[5,6,8,11–14] In future studies and implementation of TM, it is important to be aware of barriers for participation.

The aim of this study was to investigate the relationship between reasons given for refusal and the characteristics of the patients. This was done in a randomized TM trial with patients with COPD.

Materials and methods

Inclusion and exclusion criteria for the NetKOL trial (NetCOPD trial)

We recruited COPD patients with severe, but stable, COPD at high risk of exacerbations and hospital admissions.

Patients eligible for the study had to fulfil the following inclusion criteria: (1) COPD defined according to the GOLD (Global Initiative for Chronic Obstructive Lung Disease) criteria,[15] (2) post-bronchodilator $FEV_1 < 60$ per cent of predicted value, (3) hospital admission for COPD exacerbation within the previous 36 months and/or on long-term oxygen therapy (LTOT) for at least 3 months, and (4) regular scheduled visits to the respiratory outpatient clinic. Exclusion criteria were as follows: (1) an exacerbation of COPD within 3 weeks prior to enrolment requiring a change in medical treatment; (2) planned vacation or other stay outside the catchment area, for 2 weeks or more during the study period; (3) unable to participate due to language barrier or cognitive disorders; and (4) not possible to establish a working telephone line.

Patient selection

Patients were recruited from respiratory wards from four hospitals in the area of Copenhagen. Between November 2013 and April 2014, 860 patients were identified in the hospital files and subsequently screened, with 560 patients meeting the inclusion criteria and no exclusion criteria.

All eligible patients were informed about the trial by letter. This was followed up within a week with further information given over the phone about the trial objective, procedures, required computer skills, and randomization (TM vs usual care 1:1). This information given over the phone also included an introduction to the TM equipment that would be installed in their homes. During the phone call, eligible patients were asked to attend an appointment to give their informed consent. If patients expressed interest, an informed consent form was mailed, giving them the opportunity to read it at home and to discuss participation with their relatives and/or general practitioner. In most cases, informed consent was signed at the outpatient clinic, but a small group of patients with poor mobility signed the consent during a visit to their home.

Technology and service offered in NetKOL

Patients received a tablet computer with a web camera, a microphone, and measurement equipment (spirometer, pulse oximeter, and bathroom scale). Besides, patients reported changes in dyspnoea and sputum. These observations were transferred to a call centre, which were open weekdays between 9 a.m. and 3 p.m. In case of 'alarming' measurements, the patient was contacted by a respiratory nurse. Patients could also contact the call centre by phone. Video consultations with spirometry were taken regularly and as needed. For further details, see Ringbaek et al.[7]

Reason for refusing participation in NetKOL

Patients who declined participation were asked the reasons for refusal. The main reason for declining participation was registered by the interviewing nurse among the following four categories: (1) technical concerns (e.g. not wanting computer in their home); (2) personal concerns (e.g. not having enough energy due to illness, comorbidities, taking care of a sick spouse, lacking time, not feeling ill enough, or feeling too old to take in new technology); (3) wishing to continue usual care, that is, regular visits at the outpatient clinic (wishing to meet with clinical staff in person); and (4) not wishing to participate in a clinical trial (randomization procedure or completing several questionnaires). There are no established categories for refusal and therefore the four categories were chosen inspired by Gorst et al.,[10] Mair et al.,[12] and Sanders et al.[17]

Some patients gave more than one reason for declining participation but were asked to state the main reason. Only the main reason was included in the present analysis.

Patients' characteristics

From the patients' hospital file, we retrieved information on body mass index (BMI), FEV_1% predicted, The Medical Research Council (MRC) dyspnoea score (scale 1–5),[16] smoking history, and LTOT.

Data on the following were retrieved from the Danish National Registry of Patients and the Danish National Registry of Deaths: selected comorbidities (cardiovascular diseases, diabetes mellitus, skeletal/locomotion problems, neoplastic diseases, depression, or osteoporosis), hospital admissions (all-cause), and vital status during the 6-month study period.

This study has been approved by the Danish Data Protection Agency (j.nr. 2007-58-0006) and Danish Health and Medicine Authority.

Statistics

Data were analysed with the statistical package SPSS version 19.0 (IBM SPSS Statistics). The chi-square, two-sample t-tests and Mann–Whitney U tests were used, as appropriate, to compare groups of interest. Data were further analysed by logistic regression analysis, where covariates (confounders) with a p-value <0.05 in the univariate analysis were subsequently included into a multivariate analysis: gender, BMI (continuous variable), FEV_1% predicted value (continuous variable), age (continuous variable), osteoporosis diagnosed (yes or no), hospital admission for COPD exacerbation in the year prior to being invited to participate in the trial (the estimated coefficients for one, two, and at least three admissions were equal and therefore analysed as yes or no). Assumption of linearity was assessed by categorizing the variable into multiple dichotomous variables of equal units on the variable's scale. The estimated coefficients of each dichotomous variable were compared. Kaplan–Meier analysis with log-rank statistic was applied to estimate a model for time to death. A two-sided p-value of <0.05 was considered significant.

Results

Number and reasons for declining to participate

Of the 560 eligible patients, 279 (49.8%) patients declined to participate in the NetKOL trial (Figure 1). ‘Personal reasons’ were the most frequently stated reason for declining given by 164 patients (61.0%). Out of these 164 patients, 37 (22.5%) believed that weekly monitoring was unnecessary because they did not feel sick enough. In addition to ‘personal reasons’, ‘technical concerns’ were given by 55 (19.0%), ‘wishing to continue usual care at the outpatient clinic’ were given by 17 (6.3%), and ‘not wanting to participate in a clinical trial’ were given by 23 (8.6%) (Figure 1).

Characteristics of patients who declined to participate in the NetKOL trial

Compared to patients consenting to participate, patients declining participation in the TM trial were older, more often female, had a higher FEV_1%, had a lower BMI, were more frequently diagnosed with osteoporosis, and had more hospital admissions due to COPD exacerbation the year prior to the inclusion (Table 1).

The multivariate logistic regression analysis revealed that lower BMI was not independently associated with declining to participate in the trial. Female gender, higher age, concomitant osteoporosis, and previous hospital admission for COPD exacerbation were independently associated with declining to participate in the tele-healthcare trial (Table 2).

Discussion

This study shows that a significant proportion of the patients who were invited to participate in the NetKOL TM trial refused because of personal reasons and concerns regarding the use of the TM technology. More females than males refused to participate and the average age was higher.

When planning larger controlled TM trials, it should be taken into account that a large proportion of the screened patients are not likely to be eligible for the study and that the rates of refusal and drop-out are often considerable. Similar to other studies, we had to screen approximately three times as many patients as required for completing the trial[5,6,8,12] and half of our eligible patients declined participation. Other similar studies have reported a refusal rate from 22 to 80 per cent with the highest rate found among patients with acute exacerbations in COPD.[5,6,8,11–14] This supports our data, where the group of non-participants had higher rate of hospitalization prior to the trial than patients enrolled. COPD and osteoporosis are strongly associated because of common risk factors such as age, smoking, treatment with corticosteroids, and inactivity. Non-participants are characterized by having more severe COPD with more frequent hospitalization due to exacerbation in COPD. These patients are more inactive and often treated with corticosteroids which increase the risk of osteoporosis. They are less stable than the enrolled patients.

So far, only limited evidence has been published for the characteristics of patients declining participation in TM trials. In line with our findings, Mair et al.[12] showed that non-participants were older, but also that they more often than patients enrolled had inhaled corticosteroids prescribed. In our study, we unfortunately lack information on prescribed medication, but prior to the assessment of eligibility, patients declining participation had a higher rate of hospitalization caused by acute exacerbation of COPD. Since treatment is in accordance with the GOLD strategy document,[15] the majority of these patients are likely to be prescribed inhaled corticosteroids, thus supporting Mair et al.[12] Similar to our study, Jakobsen et al.[14] found that there were more females represented in the group of non-participants. One theory could be that females offered TM had less experience with

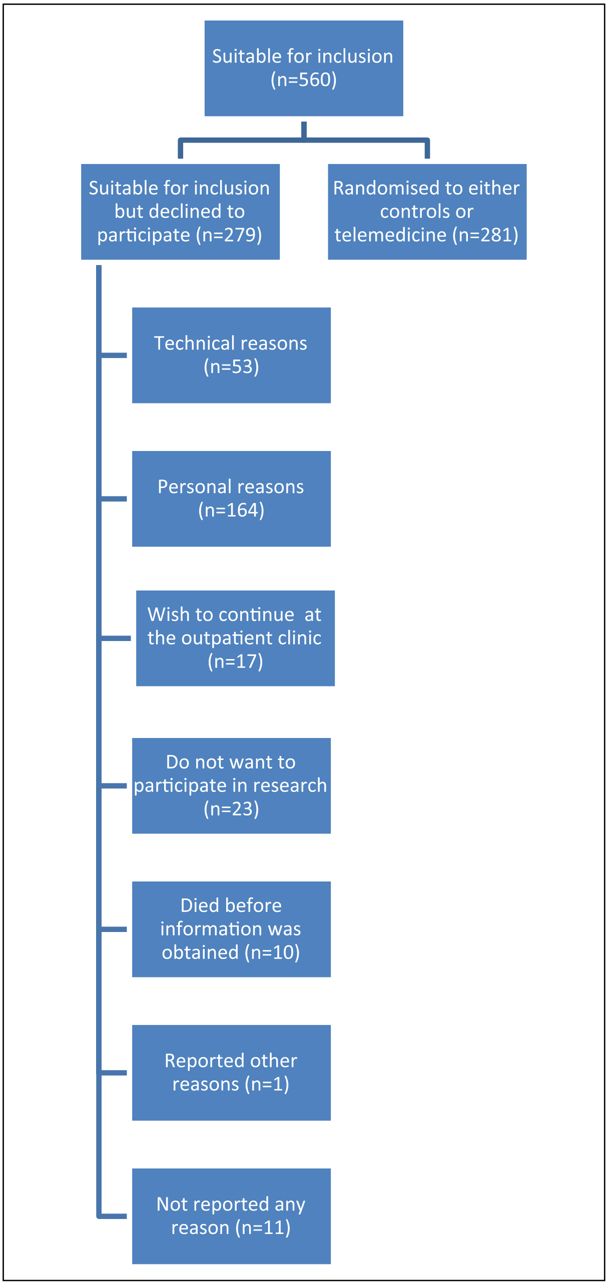

Figure 1. Patients' reasons for declining participation.

computers and technique compared to males. Unfortunately, we do not have any data on experience with computer or Internet in the group of decliners, but among those who participated in NetKOL, 50 per cent of the males answered that they often used a computer, 24 per cent never, and 26 per cent 'now and then'. The same figures for females were 38, 31, and 31 per cent, respectively.

Table 1. Characteristics of the eligible COPD patients (n = 538) eligible for inclusion in a 6-month controlled trial of Tele Healthcare divided according to whether they wanted to participate or refuse.

	Participated, N = 281	Refused, N = 257[a]	p-value
Females, N (%)	148 (52.7)	177 (68.9)	<0.001
Age (years)	69.6 (9.6)	74.0 (9.0)	<0.001
BMI (kg/m^2)	25.9 (6.9)	24.2 (5.7)	<0.001
FEV_1% predicted	34.4 (12.6)	36.8 (11.1)	0.032
MRC dyspnoea score, mean (range)	3.6 (1–5)	3.7 (2–5)	0.11
Current smokers, N (%)	82/281 (29.2)	68/250 (27.6)	0.61
Long-term oxygen therapy (%)	26.7	22.8	0.32
Cardiovascular diseases (%)	63.0	65.0	0.63
Diabetes mellitus (%)	16.0	15.2	0.79
Osteoporosis (%)	18.9	33.4	<0.001
Skeletal/locomotion problems (%)	16.4	22.2	0.09
Neoplastic disease (%)	5.0	5.1	0.97
Depression (%)	9.6	14.4	0.09
Time to death, mean (days)	847 (816–879)	785 (732–839)	0.14
Exacerbation of COPD requiring hospitalization, mean (range)	0.55 (0–5)	0.74 (0–8)	0.21
Hospital admissions for COPD exacerbation in the year prior to being invited to the trial, mean (range)	1.07 (0–23)	1.41 (0–14)	<0.001

COPD: chronic obstructive pulmonary disease; BMI: body mass index; FEV_1: forced expiratory volume in 1 s; MRC: Medical Research Council.
Values are mean (standard deviation (SD)) unless otherwise specified.
[a]A total of 22 patients were excluded from the analysis, as they did not give consent to collection of clinical data.

Table 2. The likelihood (odds ratio) for patients with COPD to decline participation in a 6-month controlled tele-health trial (NetKOL).

		Univariate analysis	Multivariate analysis
Gender	Male	1	1
	Female	1.99 (1.40–2.83)	2.01 (1.34–3.03)
Age	1 year	1.06 (1.03–1.09)	1.06 (1.03–1.09)
Osteoporosis	No	1	1
	Yes	2.17 (1.45–3.23)	1.92 (1.23–3.03)
Hospital admissions for COPD exacerbation in the year prior to being invited to the trial	0	1	1
	⩾1	2.44 (1.69–3.45)	2.04 (1.39–3.03)
Body mass index	1 kg/m^2	0.95 (0.93–0.98)	NS
FEV_1% pred	1%	1.02 (1.00–1.03)	NS

COPD: chronic obstructive pulmonary disease; NS: not significant; FEV_1% pred: forced expiratory volume in 1 s as percentage of predicted value.
The results of logistic regression analyses are given in terms of estimated odds ratios with corresponding 95 per cent confidence intervals (CIs). Variables included in the analysis are listed in Table 1.

Supporting these data, a survey from 'Danmarks Statistik' in 2014 showed that among 65- to 89-year-old Danish people, 60 per cent of males and 45 per cent of females use Internet daily.

As mentioned, personal reasons were the main reason stated for declining participation in our study. Despite the fact that all eligible patients suffered from severe COPD and had had previous hospitalizations due to exacerbation in COPD, a high figure of 13 per cent of the patients did not feel sick enough to participate, and moreover, they believed that weekly monitoring was unnecessary. When their disease was in the stable phase, these patients wanted their lives to remain normal with as little focus on their COPD as possible. Similar findings have been reported from the Whole System Demonstrator, where 19 patients declined participation because they saw it as a threat to their identity.[17] They regarded telehealth as being associated with a high degree of dependency and poor health. One patient expressed, 'I'd feel more crippled', and another said, 'I would be more stuck inside the house', 'Am I that ill – I do not need to be reminded every day of my illness. Give it to somebody who really needs it'.[17]

Technical concerns are another common barrier for TM.[10] For our patients, previous experience with computers had positive impact on completing the NetKOL trial. But this is by no means prerequisite since 30 out of 35 (85.7%) patients, who had no computer experience, completed the trial (data not shown). Unfortunately, we do not have information on the level of experience with computer among patients refusing participation in TM. Another limitation of our study is that only the main reason for non-participation was included in the analysis. This might blur or simplify a more complex constellation of underlying reasons.

As half of our patients declined to participate in NetKOL, it is worth asking if better information for the patients might have increased the participation rate. Patients were initially contacted by phone. This is probably not the ideal media for describing the technical aspects of a TM trial or to raise interest for participation among patients with generally limited computer experience. The TM equipment used in this study was not shown to the patients until after the patients had consented to participation. Declining due to technological concerns is therefore most likely based on the individual patient's expectations or presumptions. Furthermore, a quiet setting and dedicated time is essential for a successful communication of benefits of participation and for the demonstration of TM equipment. We therefore think that the participation rate might have been higher if information about the study and the actual demonstration of the equipment had been given at the same time, preferably when patient was seen at the outpatient clinic. Moreover, the health professionals did not have any special training in instructing the patient in the TM equipment and had little experience in using the equipment at the time of recruitment. We believe that in some cases, if the quality of information was not optimal, it could influence the patient's decision to participate or not. This has not been analysed in this study and requires other studies investigating the influence and need of the skills of health professionals instructing patients in TM.

Also, exploring the patients' personal needs and expectations will enable the health professional to better address the patients' concerns. In our study and other randomized clinical trials with TM, individual adjustments to the technical setup have been limited, for example, the patients did not have an option to choose between different types of equipment. One could speculate that offering the patients a choice between different types of equipment or giving the staff an opportunity to tailor the TM to the patient's needs would address some of their technical concerns.

Having said that, far from all barriers to participate due to personal concerns can be overcome by training of health professionals and thus providing higher quality of information for the patient. The patient's decision not to participate should always be acknowledged and respected.

A minority of our patients declined participation on the ground that they wished to continue regular follow-up at the outpatient clinic. All the participants lived relatively close to the hospital, so it is possible that more patients would have preferred TM, had the distances been greater. However, whether distance and longer journeys affect participation in TM has, to our knowledge, not yet been investigated.

Factors influencing the acceptance of new technologies among elderly people (>65 years old) seem to be complex, and perhaps we as respiratory healthcare providers have to accept that we still have a long way to go in understanding the complexity of our patient's lives and decision making.[18,19] Also, we should acknowledge that not all barriers to participate in TM, technical as well as personal, should be overcome. The aim is to increase recruitment of the right patient for the right treatment, and TM is not right for all patients at all times.

Conclusion

We found that about half of the eligible patients declined to participate in a 6-month trial of TM. They were more often females and representing the group of unstable patients with frequent admission to hospital with COPD exacerbation. They most frequently stated 'personal reasons' and 'technical concerns' as the reasons for refusal.

Acknowledgements

The authors wish to thank the NetKOL group (Pia Andersen, Lisbeth Østergaard, Christine Lilliedahl, Torben Lage Frandsen, Jo-Ann Ramsrud Jensen, Lene Nissen, Marie-Louise Pagh Søndberg, Tanja S Hansen, and Zofia Mikolaczyk) for data collecting and for their enthusiastic collaboration throughout the study period.

Declaration of conflicting interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) received no financial support for the research, authorship, and/or publication of this article.

References

1. Wootton R. Twenty years of telemedicine in chronic disease management: an evidence synthesis. *J Telemed Telecare* 2012; 18(4): 211–220.
2. De San Miguel K, Smith J and Lewin G. Telehealth remote monitoring for community-dwelling older adults with chronic obstructive pulmonary disease. *Telemed J E Health* 2013; 19(9): 652–657.
3. Dinesen B, Haesum LK, Soerensen N, et al. Using preventive home monitoring to reduce hospital admission rates and reduce costs: a case study of telehealth among chronic obstructive pulmonary disease patients. *J Telemed Telecare* 2012; 18(4): 221–225.
4. Vitacca M, Bianchi L, Guerra A, et al. Tele-assistance in chronic respiratory failure patients: a randomised clinical trial. *Eur Respir J* 2009; 33(2): 411–418.
5. Steventon A, Bardsley M, Billings J, et al. Effect of telehealth on use of secondary care and mortality: findings from the Whole System Demonstrator cluster randomised trial. *BMJ* 2012; 344: e3874.
6. Pinnock H, Hanley J, McCloughan L, et al. Effectiveness of telemonitoring integrated into existing clinical services on hospital admission for exacerbation of chronic obstructive pulmonary disease: researcher blind, multicentre, randomised controlled trial. *BMJ* 2013; 347: f6070.
7. Ringbaek T, Green A, Laursen LC, et al. Effect of tele health care on exacerbations and hospital admissions in patients with chronic obstructive pulmonary disease: a randomized clinical trial. *Int J Chron Obstruct Pulmon Dis* 2015; 10: 1801–1808.
8. Jódar-Sánchez F, Ortega F, Parra C, et al. Implementation of a telehealth programme for patients with severe chronic obstructive pulmonary disease treated with long-term oxygen therapy. *J Telemed Telecare* 2013; 19(1): 11–17.
9. Sorknaes AD, Madsen H, Hallas J, et al. Nurse tele-consultations with discharged COPD patients reduce early readmissions – an interventional study. *Clin Respir J* 2011; 5: 26–34.

10. Gorst SL, Armitage CJ, Brownsell S, et al. Home telehealth uptake and continued use among heart failure and chronic obstructive pulmonary disease patients: a systematic review. *Ann Behav Med* 2014; 48: 323–336.
11. Lewis KE, Annandale JA, Warm DL, et al. Does home telemonitoring after pulmonary rehabilitation reduce healthcare use in optimized COPD? A pilot randomized trial. *J Telemed Telecare* 2010; 16(5): 253–259.
12. Mair FS, Goldstein P, Shiels C, et al. Recruitment difficulties in a home telecare trial. *J Telemed Telecare* 2006; 12(Suppl. 1): 26–28.
13. Antoniades NC, Rochford PD, Pretto JJ, et al. Pilot study of remote telemonitoring in COPD. *Telemed J E Health* 2012; 18(8): 634–640.
14. Jakobsen AS, Laursen LC, Rydahl-Hansen S, et al. Home-based telehealth hospitalization for exacerbation of chronic obstructive pulmonary disease: findings from 'the virtual hospital' trial. *Telemed J E Health* 2015; 21(5): 364–373.
15. Vestbo J, Hurd SS, Agusti AG, et al. Global strategy for the diagnosis, management, and prevention of chronic obstructive pulmonary disease: GOLD executive summary. *Am J Respir Crit Care Med* 2013; 187(4): 347–365.
16. Fletcher CM, Elmes PC, Fairbairn AS, et al. Significance of respiratory symptoms and the diagnosis of chronic bronchitis in a working population. *Br Med J* 1959; 2: 257–266.
17. Sanders C, Rogers A, Bowen R, et al. Exploring barriers to participation and adoption of telehealth and telecare within the Whole System Demonstrator trial: a qualitative study. *BMC Health Serv Res* 2012; 12: 220.
18. Peek S, Wouters EJM, van Hoof J, et al. Factors influencing acceptance of technology for aging in place: a systematic review. *Int J Med Inform* 2014; 83; 235–248.
19. Joyce K and Loe M. A sociological approach to ageing, technology and health. *Sociol Health Illn* 2010; 32(2): 171–180.