Aalborg Universitet

AALBORG UNIVERSITY
DENMARK

Proceedings from The 16th Scandinavian Conference on Health Informatics 2018, Aalborg, Denmark August 28–29, 2018

Bygholm, Ann; Pape-Haugaard, Louise; Niss, Karsten Ulrik; Hejlesen, Ole; Zhou, Chunfang

Publication date:
2018

Document Version
Publisher's PDF, also known as Version of record

[Link to publication from Aalborg University](#)

Proceedings from
The 16th Scandinavian Conference on Health Informatics

2018

Aalborg, Denmark August 28–29, 2018

Editors

Ann Bygholm, Louise Pape-Haugaard, Karsten Niss, Ole Hejlesen, Chunfang Zhou

Scientific Program Committee

Pernille Bertelsen, Denmark
Andrius Budrionis, Norway
Ann Bygholm, Denmark
Mariann Fossum, Norway
Gunnar Hartvigsen, Norway
Ole Hejlesen, Denmark
Louise Pape-Haugaard, Denmark
Anne Marie Kanstrup, Denmark
Carl E Moe, Norway
Karsten Niss, Denmark
Elin Thygesen, Norway
Vivian Vimarlund, Sweden

Sponsors
DaChi – Danish Centre for Health Informatics
Aalborg University

Table of Contents

Papers

Abstracts

Three Living Labs in Denmark:
Challenges with Co-design and Implementation of Health IT

Tariq Osman Andersen[a], Anne Marie Kanstrup[b], and Signe Louise Yndigegn[c]

[a]*Department of Computer Science, University of Copenhagen, Copenhagen, Denmark*
[b] *Department of Planning, Aalborg University, Aalborg, Denmark*
[c]*IT University Copenhagen, Copenhagen, Denmark*

Abstract

Living labs are increasingly used as an approach for facilitating innovation and testing emerging information technologies. In this paper we analyse three large-scale technology design projects in Danish healthcare where co-design and implementation activities were organised in living labs. We describe some of the critical challenges that we experienced from transitioning technology prototypes and co-design activities into becoming part of the daily lives of patients, citizens and healthcare practitioners. The main challenges relate to creating and sustaining new work practices, scaling the number of participants, and facilitating the transition between everyday life and living lab behaviour.

Introduction

Information Technology (IT) offers continuously new opportunities for supporting health and has been pointed out as a driver for healthcare innovation [1]. However, the socio-technical aspects of healthcare systems are complex, and research stresses the importance of engaging key actors in the co-design and implementation of future technology to bridge knowledge gaps in requirements analysis and design for functional support and quality in use [2]. Living laboratories, in short living labs, is an increasingly used approach for facilitating co-design of emerging health technology across multiple actors including engineers, researchers and end-users. Living labs are technological installations set-up in (semi) naturalistic settings over a medium- or long-term period providing an infrastructure for transitioning design to implementation. As such, living labs complement usability labs, which have been proven successful for quick findings of many usability issues but often miss aspects related to in-situ use of the technology [3; 4]. Living labs offer a way for exploring and continuing design along with implementation in cooperation with future users [5]. This focus on continued co-design goes hand in hand with a new wave of studies and critical voices that suggest considering 'success' in technology co-design projects to be more focused on the post-project impact among users and participating stakeholders like software companies, funding agencies and the wider society [6]. However, integrating co-design of future health technology in living environments for health care is a complex intervention and a fundamental challenge for the design and innovation community. Yet, as highlighted by Bygholm and Kanstrup [7] the living lab literature is in general characterised by prescriptive examples with limited insights and reflections on challenges for setting up and managing living labs to meet the intentions of long-term participatory technology innovation. Examples of reported challenges are motivation of users during the lab period [8], facilitating innovations among various stakeholders [9], and the importance of maturing the technical set-up [8].

In this paper we analyse some of the major challenges that we have encountered first-hand as researchers in three large-scale health IT design projects where co-design activities were organised in living labs. The remaining of this paper provides first, background on co-design in living environments. Second, we present the three living labs and their challenges. Third, we conclude on core challenges across the three cases and consider perspectives for research in co-design and implementation in living health environments.

Co-design in Living Environments

The attention on moving design activities into use settings to bridge the gap between technology innovation and use practice is not entirely new but found as a key methodologically focus in studies that aim at diffusion and adoption of innovation [10, 11]. However, as presented by [12] co-design research is predominantly focused on early stages of requirements analysis also known as 'the fuzzy front end of design' because of the initial unclear conceptual understanding of the product being developed. Living labs offer an attention for organising long-term participation around design, development, implementation and use. However, the approach is unclear, or diverse. As stated by Bannon and Ehn [13], examinations of what actually goes on in living labs are scarce.

Hence, research on co-design and implementation in living labs is relevant.

The use of the term living lab is broad and makes the concept difficult to grasp. Some living labs aim to be real world test beds setting up a "wireless playground" for "ordinary people on the street" like in the European OPIUM networks of living labs [14] or in small rural villages like Wray Living Lab [15]. Other living labs are controlled environments set-up in test centres furnished like a "real home" and inhabited by volunteer participants in the AwareHome at Georgia Tech [16] and the PlaceLab at MIT [17]. Yet other labs are a mix of naturalistic environments and controlled set-ups in family homes like the SMEDL lab [8] and local stores in a city like Living Lab Skagen [18]. Methodological concerns in the living lab literature are often related to the dilemmas of how artificial to make the naturalistic setting – how to find the balance between controlled research and keeping sense of the real-life practice [8]. Open innovation is an often-used term [9, 19] presenting the ideals of living labs as platforms facilitating innovative co-operation among stakeholders. In the naturalistic end of the scale, labs present the aim of making an e-Infrastructure that facilitate citizens' participation as the central objective and challenge of living labs [14]. In contrast, artificial labs present the opportunity for detailed documentation and systematic test of technological installations as the core objective [16, 17]. Mixed labs are occupied with setting the stage for co-operations among designers and users [8, 18]. Data collection in living labs often includes a mix of unobtrusive and obtrusive methods like data-logs, observations, users' self-documentation, interviews, and workshops [20]. Only few living lab studies goes at length in describing how the innovation is facilitated and the challenges that arise with regards to e.g. long-term user involvement [8].

Three Living Labs in Danish Healthcare

To detail our understanding of co-design and implementation in living labs for future health IT we analyse three living labs with attention on identifying core challenges within and across the cases.

Lab 1: Technology supported senior networks

In a first living lab, the aim is to support and create new networks between senior citizens based on ideas of sharing and helping each other. The living lab is part of a project Give&Take (2014-2017) (http://givetake.eu) developed in the framework of a European project with partners in Denmark, Portugal and Austria. The aim was to co-design digital mediated sharing within senior communities (e.g. IT-volunteers from the library, or a walking group). The platform developed in the project allows senior citizens to reciprocally exchange services and resources. The platform is designed to support existing and often loosely coupled communities in order to strengthen and sustain. At the same time the idea is that the communities through the platform stays connected to a 'coordinator' like a health counsellor, social worker or the like who initiated or helped establish the community, who can then remotely follow the group (after the

initiation) on the platform and here inspire the group with other offers and possible activities - and only reach out with a helping hand when needed. The platform is for vulnerable communities that need support for networking. It was developed during the project's first year, based on the outcome of a series of dialogue meetings and later workshops. Here 50-60 senior citizens and employees working with senior communities explored together with researchers, the municipality and private partners what sharing is in senior communities and how it could be supported by a digital platform.

Around nine small living labs were established after the co-design workshops, where a finished trial version of the platform was tried out to explore whether and how it could support and optimise sharing and exchange among members of a senior community. The Living labs were established by creating arrangement with different senior communities and the connected institutions (municipality) or organisations (e.g. DanAge). It was communities like a walking group, a group of IT volunteers, and a food club for men. The third living lab lasted throughout the last part of the project, while the rest only ran for a couple of months. Researchers and municipality representatives took part in the communities' activities – in some of the communities on a weekly basis for 4-5 months – to observe the communities as well as to introduce the technology and create small experiments for the exploration of the technology. The living labs became a space for rehearsing new practices both for the citizens in the communities and for the 'coordinators' from the municipalities, etc. [21] – and the aim was to make this practice viable and sustainable after the end of the project. This became challenging especially in relation to the coordinators.

One thing was to try out and evaluate the digital platform in the living lab, but the rehearsals of new practices turned out to be very challenging especially in the transition from 'rehearsing new practices' to 'make the new practices viable and sustainable' beyond the project and the engagement of the researchers. The researchers realised that their participation in the community was not just researchers intervening and observing the group's use of the platform. Their participation became a rehearsing of the coordinator's new practices. The introduction to and support with the platform - and the interaction with the community through the platform were all different ways of trying out what the coordinator role could be like and what would be required of the coordinator in relation to creating a community supported by a digital sharing platform. Though, the researchers' number of hours of presence in the meetings, home visits and on the platform was not realistic and ideal for a coordinator. Especially due to the idea that the digital platform should support self-organising groups – coordinators were only meant to remotely follow the group and only reach out with a helping hand when needed.

The question the project wanted to explore was how the platform could be valuable in the 'coordinators' daily work? What kind of support and what kind of functions on the platform and information (or data) would they need? Though, the initial part of introducing the platform, etc. took a great amount of time. It was necessary for the implementation, but it made difficult to explore these questions. After the project

2

and the researchers left the living labs, 2-3 of the communities continued using the platform. In one of the cases, the coordinator also continued to a lesser degree. Without the framing of project, it was no longer an articulated part of the coordinators' work. At the same time, the platform was not robust enough to be sold to the municipalities, social housing associations, etc. - it required further development. The platform is therefore not available to new communities (if the coordinators would like to share it), while at the same time the existing communities cannot get IT support from either a 'Give&Take team' or the coordinators. The aim of creating viable and sustainable practices becomes difficult especially when the mediating platform is still unstable and the organisations are not committed to it (after the project), which raises question on how to navigate in these living labs with new technology and practices when the aim is to support social aspects (to create new or stronger communities), if one or more of the actors that take part in mediating the social practice are uncertain or unstable.

Summing up, central challenges in the Living Lab 1 was:

- Challenges with transition: The process from rehearsing to real-life practice was difficult and unclear since the living lab is only a minor part of the coordinators work and the practices were not merged with or adapted to the existing practices of e.g. the Health centre.

- Challenges with resources: Lack of time and resources after the project ended made it difficult to sustain the new digital (and social) practices.

- Challenges with commitment: Commitment from all partners to engage in the temporary practice in the Living Lab was a challenge. The ambition that citizens adapt to the new practice is difficult when the other partners only commit to the practice within the timeframe of the living lab.

Lab 2: e-Health for Heart Patients

In a second living lab, the aim is to improve remote communication between patients with an advanced pacemaker and clinicians. The living lab is part of a large R&D project called SCAUT (http://scaut.dk) that runs from 2014-2018, supported by the Innovation Fund Denmark (#72-2014-1) with a consortium of two industry and two public partners including The Heart Centre at Rigshospitalet in Copenhagen Denmark. The project began with IT design researchers doing fieldwork studies and co-design activities in the clinic and in patients' homes. Six months into the project, a first prototype of a patient mobile app and a web application for clinicians was developed and deployed among 20 cardiac device patients who were invited to take part in the living lab. The patients could use the app to record symptoms and describe their context in audio while clinicians at the Heart Centre could use the patient-generated data for decision-making and for providing feedback to patients. Satisfaction was quite high among patients, since it supported better contact with the clinic than without the app. Many patients explained that it made them feel safer and more informed. The

clinicians, on the other hand, were at times dissatisfied due to some features, such as a symptom diary, introduced more work [22]. Nurses and bio technicians explained that although the system supported provision of better care it also generated new accountabilities and tasks that there was no time for. And since the prototype was only used in follow-ups with 20 patients, and not the 3.500 cardiac device patients at the Heart Centre, it became a general concern that the system could not scale. For the prototype to work, the design researchers now had to focus on re-designing the functionality so that it provided value for patients and clinicians at the same time [22]. The question was therefore: How to adapt the prototype so that it could become useful for the actual, large-scale remote monitoring work? To answer this question, the design researchers decided to do two things: Increase the efforts in re-designing the features that were not adding value in the clinic and increase the number of participating patients to ensure evaluation against real-life, large-scale remote monitoring work.

Over the course of two years, the design researchers succeeded to onboard more than 200 patients and a total of nine clinicians as well as adapting the prototype to become more useful for clinicians. The scaling up of user involvement and co-design in the living lab was, although, a very difficult undertaking: Small technical issues for a few users are now large critical issues for many users; A few good ideas and design inputs for changes in the prototype are now hundreds of ideas and inputs for adaption; High user engagement is now high, medium or no user engagement; Good personal relations with a few users is now good, little or no personal contact with many users.

Technical/practical issues as well as design issues multiply and increase when scaling up. For example, it became a time-consuming task in itself to support patients in understanding and using the prototype. Other important tasks that arose was keeping track of and reporting back to participants as well as coordinating when and how to contact them. The ability to remember and differentiate among participants, their expectations, and the degree to which they wish to engage became an issue. More time was needed to monitor use of the system along with reaching out over the telephone to users and non-users to learn about the reasons and discuss ways for improvement.

Summing up, central challenges in the Living Lab 2 was:

- Challenges with scaling: The co-design with 200 users was difficult and identified a need for resources and approaches for co-design in large scale living labs.

- Challenges with technical issues: Small technical problems for a few users became large-scale technical issues for the 200 users and required strong focus on a mature and reliable prototype.

- Challenges with practical issues: Practical tasks like coordinating user communication, keeping track of users and the use etc. became a central (underestimated) task in the living lab.

Lab 3: Welfare Technology in Nursing Homes

LabX was the overall name for an umbrella of living labs set up in North Denmark to explore living labs for innovation of healthcare technologies in the care of elderly, chronically ill and handicapped persons. The project was supported by the European Regional Development Fund and the overall goal was to foster collaboration between municipalities, industry and knowledge institutions in order to stimulate economic growth for the industrial partners, achieve efficiency and cost reduction in the public sector and better serve the citizen. Six municipalities and eleven small- and medium-sized technology enterprises participated in the project together with a number of nursing homes, a university college, a vocational education, and researchers from Aalborg University. Thirteen living labs were initiated but only eight living labs were succeeded on a basis where data quality allowed for analysis. A variety of technologies were explored. Some technologies were aimed at staff only, but most technology were intended to support both residents and staff. Technology included digital fences based on sensors and Bluetooth bracelets to sound alarms of residents left the nursing home. Bluetooth bracelets with accelerometers to send alarms if residents have a fall. Sensor screens to stimulate residents via digital art. Automatic toilets, intelligent beds, intelligent laundry based on RFID chips, a machine to help residents into and out of compression stockings. Seven of the eight nursing homes were intended for elderly people and one was for young adults with physical disabilities. In all cases living labs were set up for a three-month period beginning with a contract between participating nursing homes and technology enterprises regarding the purpose of the lab.

The living labs started with the technical installation and short training of super-users, i.e. selected staff trained to use the technology and record data. A living lab coordinator visited the labs on a frequent basis to collect data and support the coordination of activity. However, the initial idea was to install the technology and then observe the use. This approach resulted in a low use of co-design methods facilitating interaction among the various stakeholders in the living labs. Instead data collection was based on individual interviews and observations with staff, management, municipality and residents. Only in one lab a co-design workshop was organised. The participation of the nursing home residents was limited.

The ambition to set-up a living lab and then observe the lived life with the new technology was based on the assumption that a living lab (in contrast to a simulation or usability lab) is close to a naturalistic environment because of the long-term installation and use in a living environment. However, though the labs lasted for three months the results repeat existing concerns from the living lab literature on missing long-term perspectives – start-up-problems continued throughout the labs with only one exception. An example is the living lab with the intelligent beds, which started with a massive amount of alarms send from the beds. Alarms make a high sound in the hallways. As described by the care workers in interviews this was 'very disturbing' and the beds were all turned off within the first 24 hours. After a couple of weeks where the producer worked with the technical installation the beds were re-installed and used for three months. However, problems with false alarms and missing alarms continued due to a weak technical infrastructure between the beds and the nursing home's existing network. Some alarms reached the mobile phones of the care workers, and some alarms did not. The technical problems moved attention from daily living in the lab. As expressed by the manager: 'It is frustrating because I thought that the focus was different, to use the bed in our care, and then the bed was not in focus at all, it was the alarms and the paging system that got all attention'. Similar technical problems stealing the attention from the 'real-life' was observed in other labs [23].

A consequence of assuming that living labs are close to real life living is the assumption that life in living labs is business as usual – technology can be installed and life can continue as usual and likewise, technology can be removed after three months and life can continue as before. Assuming business as usual means to assume that people who live and work in labs know what to do, i.e. that they are familiar with the environment (since it is their daily environment) and know how to do their daily tasks (as usual). However, in these labs the roles and activities were mostly unclear to most of the participants. Are care assistants and residents allowed to unplug technology if it keeps firing alarms? In other words, are care workers testers who must work with and report errors? Are they innovators who must come-up with solutions and engage in technology development? Or are they simply expected to behave as care workers and residents and do business as usual? In general, the labs identified a need to define lab behaviour

Summing up, central challenges in the Living Lab 3 was:

- Challenges with participation: The interaction between the multiple stakeholders in the living labs was low and mostly non-existing and the participation of the elderly users was very limited. This identified that co-design activities need to be methodological designed as part of a living lab set-up.

- Challenges with exploring the real-life: Technical problems were stealing resources and attention throughout most of the living labs and identified a need to define and develop approaches for innovation in real-life health settings.

- Challenges with lab behaviour: unclear roles and lab behaviour caused conflicts and misunderstandings and identified a need to define roles, tasks and scrips for living lab.

Conclusion

In this paper, we have analysed and identified nine challenges within three health IT living labs in Denmark; Transition, resources, commitment, scale, technical issues, practical issues, participation, real-life and lab behaviour. Each of these challenges are interrelated. In lab 1, the transitioning from design to implementation and use was very challenging since it hinges on organisational commitment and securing re-

sources for continuing stabilisation of the IT platform after the living lab. In lab 2, large-scale user participation accentuated the technical issues and practical issues in the ability to facilitate transition from design to use –technical and practical issues affected the lab participants commitment and experience. In lab 3, the tension between real-life and lab behaviour became a challenge, primarily due to unclear roles and little interaction among stakeholders. Thus, a general conclusion from the analysis is, that examinations of living lab challenges are needed to further advance approaches for health IT innovation in living environments. An examination of the challenges across the three examined living labs in this paper indicate that the living lab approach is indeed a socio-technical challenge calling attention to the need to facilitate the complex interrelation between technology, humans, organisational structure and tasks when innovating new technology supported health practices [24].

Acknowledgments

References

[1] World Health Organization. 2005. *Connecting for Health: Global Vision, Local Insight*. Report for the World Summit on the Information Society. WHO.

[2] N. Carroll and I. Richardson. 2016. Aligning Healthcare Innovation and Software Requirements through Design Thinking. *SEHS'16.*

[3] Kjeldskov, J. et al. Is it worth the hassle? Exploring the added value of evaluating the usability of context-aware mobile systems in the field. *Mobile HCI'04*, Springer, (2004), 61-73.

[4] Jensen, S. et al. 2015. Clinical simulation: A method for development and evaluation of clinical information systems. *Journal of Biomedical* Informatics, 54: 65-76.

[5] Aanestad, M. et al. 2017. Participatory Continuing Design: "Living with" Videoconferencing in Rehabilitation. In A.M. Kanstrup et al. (Eds.) *Participatory Design & Health Information Technology*. IOS Press.

[6] Kyng M. Bridging the Gap Between Politics and Techniques: On the next practices of participatory design. *Scandinavian Journal of Informaiton Systems*. 2010; 22(1):5.

[7] Bygholm, A. & Kanstrup, A. M. 2017. This Is not Participatory Design – A. Critical Analysis of Eight Living Laboratories. In A.M. Kanstrup et al. (Eds.) *Participatory Design & Health Information Technology*. IOS Press.

[8] Ogonowski, C. et al. 2013. Designing for the Living Room: Long-Term Users Involvement in a Living Lab. *Proc CHI'13*, ACM Press: 1539-1548.

[9] Følstad, A. 2008. Living labs for innovation and development of information and communication technology: a

literature review. *eJov: The electronic journal for Virtual Organization & Networks.*

[10] Rogers, E. 2003. *Diffusion of Innovations, 5th edition*. Free Press.

[11] von Hippel. 2001. Perspective: User toolkits for innovation. The journal of Product Innovation Management, 18: 247-257.

[12] Sanders, E. and Stappers, P. J. 2008. Co-creation and the new landscapes of design. *CoDesign*, 4, 1: 5-18.

[13] Bannon, L.J. & Ehn, P. Design: design matters in Participatory Design. *Routledge International Handbook of Participatory Design,* Simonsen, J. & Robertson, T. (eds.). Routledge (2012), 37-63.

[14] Ponce de Leon, M. et al. 2006. Creating a distributed mobile networking testbed environment – through the Living Lab approach. *2nd International Conference on Testbeds and Research Infrastructures for the Development of Networks and Communities*, IEEE, 1-5.

[15] Ishmael, J. et al. 2010. Deploying rural community wireless mesh networks. *IEEE Internet Computing, 12*, 4, 22-29.

[16] Abowd, G.D. et al. 2002. The Aware Home: A living laboratory for technologies for successful aging. *Proceedings of the AAAI-02 Workshop "Automation as Caregiver", AAI Technical Report WS-02-02*.

[17] Intille, S.S. et al. 2005. A Living Laboratory for the Design and Evaluation of Ubiquitous Computing Technologies. *CHI'05 Extended Abstracts,* ACM Press, 1941-1944.

[18] Kanstrup, A.M. et al. 2010. A living laboratory exploring mobile support for everyday life with diabetes. *Wireless personal communications, 53,* 3, 395-408.

[19] Bergvall-Kåreborn, B. et al. A Milieu for Innovation – Defining Living Labs. *2nd ISPIM Innovation Symosium,* New York, 2009.

[20] Erikson, M.; Niitamo, V.P. & Kulkki, S. *State-of-the-art in utilizing Livng Labs approach to user-centric ICT innovation – a European approach*. Luleå: Center for Distance-spanning Technology. Luleå University of Technology, Sweeden, 2005.

[21] Yndigegn, S. & Aakjær, M. K. Living lab: Format for rehearsing a new (service) practice. In *Proceedings of the Participatory Innovation Conference 2017*, 111-119.

[22] Andersen, T. et al. 2017. From Prototype to Products: Making participatory design of mHealth commercially viable. In A.M. Kanstrup et al. (Eds.) *Participatory Design & Health Information Technology*. IOS Press.

[23] Kanstrup, A. M. 2017. Living in the Lab: an analysis of the work in eight living laboratories set up in care homes for technology innovation. *CoDesign*, 13,1: 49-64.

[24] Bostrom, R. et al. 1977. MIS Problems and Failures: A Socio-Tehcnial Perspective. *MIS Quarterly*, 1, 3.

Address for correspondence

Tariq Osman Andersen, tariq@di.ku.dk, Department of Computer Science, University of Copenhagen, Emil Holms Kanal 6, building 24, 5th floor, 2300 Copenhagen S, Denmark

Making Computer Games that Can Teach Children with Type 1 Diabetes in Rural Areas How to Manage Their Condition

Svein-Gunnar Johansen[a,b], Eirik Årsand[b], Gunnar Hartvigsen[b]

[a]*Department of* Computer Science, University of Tromsø – The Arctic University of Norway, Tromsø, Norway
[b]*Norwegian Centre for eHealth Research, University Hospital of North Norway, Tromsø, Norway*

Abstract

Computer games can teach children a number of skills. But in order to cultivate enough engagement so that players will want to learn, the games must be sufficiently entertaining. Making good computer games is not trivial, and also not something strictly sticking to a method or script can accomplish. In the CADMOS project, we have tried to tap into kids' general interest and fascination with computer games, to teach children aged 8-12 with Type 1 diabetes how to deal with their condition in an optimal way. This will be achieved by the use of serious games that are easy to understand, yet fun to play, where they can experiment with variable treatments of their own illness in a safe space on virtual avatars instead of themselves. We also want to achieve synergistic integration with other diabetes-related treatment and self-management tools, which are already being used by children in the target group. Furthermore, it is a goal that the children's friends and family members should also be able to participate in the game and thereby gain a better understanding of what it means to live with diabetes. In this paper we show how we can get closer to this goal by designing the game iteratively together with members of our user group.

Keywords:

Video games, Diabetes Mellitus, Self-Management.

Introduction

Computer games can - as long as they are fun to play - be a valuable tool to teach children any number of skills. After all, it is easier to learn something when you enjoy doing it, and games can be a valuable source of inspiration if you live in a rural area where access to other people is limited.

The problem is that in order to get enough engagement so that players will stick with a game, just taking the game and wrapping it around some learning material is not enough. While there is little doubt that children learn from games, there are very few games that are able to specifically teach a particular skill. In order to accomplish this, the game should first and foremost be something that the intended user would want to spend time with. Creating something that qualifies in this regard is however more of an art form than an exact sci-ence. Even in commercial game development, where you only need to create something fun and not worry about teaching, the ratio of what becomes successful is very low. But it is easier to get there by involving the intended users in the design.

The core part of the CADMOS project is the development of a serious computer game for connecting children and adolescents with T1DM (Type 1 diabetes) in rural areas. Our hypothesis is: By combining mobile phones, medical sensors, social media and serious video games, a motivational and useful educational tool can be constructed, improving the self-management skills of young T1DM patients considerably.

The project has an experimental, user-oriented approach, and includes an in-depth analysis of the problem area, including social video game design for children and adolescents. Our prototype game is based upon development experience and published research on game development and social media. The project has involved children and adolescents with T1DM and their parents [1], but also researchers and developers of diabetes technology and self-management systems.

The CADMOS project is part of on-going research in Tromsø, Norway, on serious games for children and adolescents with T1DM, and includes several computer games [2-6].

Only a few existing games for children with T1DM have been made, and even fewer are generally available. [7] Two of the most interesting are "Diabetic Dog" [8] and "Carb Counting with Lenny" [9].

The Diabetic Dog Game is a serious game from Sweden (Nobelpriskampen 2009; Nobel Web AB, 2010) [10], where the users must take care of a dog with T1DM. Blood sugar levels, insulin levels, and other parameters such as mood affect the dog, and the player must make decisions and actions accordingly. The main goal of the game is to take care of the dog and make sure it is happy and healthy by giving love and affection, arranging walks, providing food, and supplying insulin.

In 2011, Medtronic released the game "Carb Counting with Lenny" (Medtronic, 2011). The game contains four mini-

games. The goal for all four games is the same – to increase knowledge about carbohydrate content in different food groups. In this way, the children can learn to manage their own food intake. It consists of two major parts:

1. Lenny's Food Guide helps kids learn carb values for many food items across the basic food groups.

2. In Lenny's Carb Games, children can test their knowledge with four interactive games: Carb or No Carb, Compare the Carbs, Guess the Carbs, and Build a Meal.

In this paper we describe the iterative design approach used in the CADMOS project, and how we have been able to engage our user group by making them part of the process.

Materials and Methods

The development project has been through two iterations:

1. Initial development work on game mechanics suited for teaching, and getting feedback on design from fellow computer game designers.

2. Presenting the game to kids in the target age group, and receiving feedback and suggestions on how the game can be improved.

We have used an ethnographic approach to gather information on how the users experience our game. We observe them whilst they play, paying particular attention to non-verbal cues as well as what they are saying, in order to determine whether they are enjoying the experience or not. We also make notes of what parts are working as intended and what parts need more work.

Stage 1: Developing the initial prototype

The initial development started as part of "Tromsø Gamelab", a one-off experiment at UIT – The Arctic University of Norway. This collaboration between academia and local game developers was aimed at creating a curriculum combining computer science with commercial and practical aspects of designing, developing and releasing a video game.

We decided to create a battle-arena game, where you pit a team of characters against an opposing team to see who wins. The plan was to create a simple but functional game mechanic, to use as a starting point for further development into something that could teach diabetes management to children and adolescents.

In order to justify putting the diabetes related parts into the game at a later stage, we created a backstory that would facilitate this. The game is set in a distant future, where humanity is genetically and mechanically enhanced in ways that practically gives them superpowers. The downside to this enhancement is that it also gives them the functional equivalent of diabetes, and thus everyone is heavily reliant on injecting insulin.

Stage 2: Getting feedback and improving the design

The next step was to test whether we were on the right track by presenting our game to children in the appropriate age group. This was done as part of a workshop organized by members of our local diabetes community. Our audience was 11 kids aged 12-17 and their parents.

This event gave us a chance to demonstrate and talk about our project to both the children and their parents. As part of the workshop, we also invited the kids to participate in the development process by designing new characters, giving feedback on what was already implemented and coming up with ideas for how to make the game even better.

Figure 1- Participants of the game design workshop

In order to get somewhere concrete during our session, we settled on one idea to focus on: *How to visualize the balance between fullness level (with regards to food) together with blood glucose level?* Preferably in a manner that would be easy to read and understand. We then workshopped a possible implementation together with the kids, using paper and whiteboard.

Results

The initial game design went through a number of visual styles before we settled on something that appealed to the other developers. This was the design we presented to the kids:
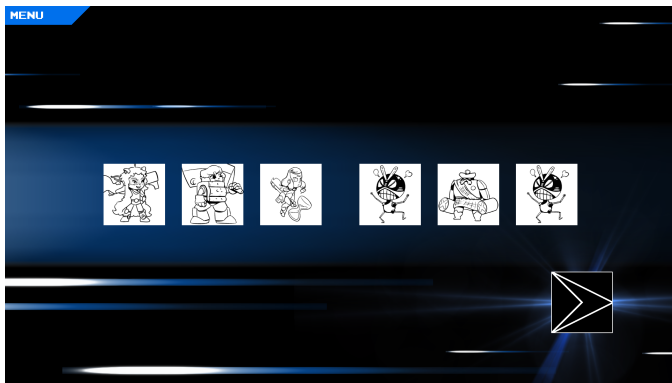
Figure 2- Screenshot from the game.

Upon doing so, we learned that the current state of our game appeals more to boys than to girls. According to one of the participants it also appeals even more to those who like science fiction in particular.

About a third of the kids really liked being involved in the design process. Another third was somewhat passive, and the last third was actively disinterested. Practically all those that liked being part of the process volunteered to continue working with the game as testers after the end of the workshop.

The workshop resulted in a prototype designs for important user interface elements that can be used in the game. Since the intended users of the product helped design it they should also be able to more easily make use of it.

Discussion

Part of the reason we chose to make the characters into cyborgs with diabetes is that an earlier workshop with the children from when they were younger, indicated that humanoid characters were usually envisioned when they were asked do draw up suggestions for games, and thus the most likely avatar type to elicit engagement [11].

Figure 3- Sample drawings from a similar workshop in 2013

But the cyborg characters also make it easier for us to handle things like death and injuries in game. We wanted the option to bring back to life characters that were seriously injured or killed, and the robotic aspect makes it plausible that characters can be "repaired".

The main reason for this buffer against fatal consequences is that we want to encourage experimentation, as that is one of the best ways to learn. In real life however, people with T1DM who are dependent on manual insulin injections can potentially die from complications associated with incorrect dosages. Badly managed insulin and blood glucose levels over time can also lead to disabilities like blindness and kidney failure. It is therefore not advisable to do experimentation with one's own body, but a computer game provides an arena that allows it to be done on avatars in a safe space.

By allowing the characters in the game to be repaired should anything go wrong we could also keep any emotional bonds the players have developed to them intact.

As the primary purpose of the work so far is putting together a game that children like playing, it is currently difficult to determine whether they are actually learning anything from it. This will eventually be something we have to test by comparing children with T1DM who play the game, with a control group of children with the same illness, but no access to the game. If we find that blood glucose levels are closer to the ideal in the first group after an appropriate amount of exposure to the game than in the second, we can conclude that it is likely working as intended.

The next phase of the project will be to import health data from on- and off-body monitoring equipment into the game. The idea is that data from sensor equipment such as step counters, glucose meters, digital body thermometers, etc., can be integrated as part of the experience.

We plan on several extensions. One idea is to let the player take the role of a diabetes adviser, who assists patients with their day-to-day activities. Each patient will present a different situation/problem that they need help with. The player will be able to see recent blood glucose measurements, dietary information, physical activity, and to ask the patient questions, and based on this, give advice. Based on the actions performed, the player will be rewarded points and achievements and the virtual patient will be either happy or unsatisfied with the help they received. The points and achievements received can be posted to an online leader board, and to the players social media profile, thus making it sharable with other people also playing.

We also want to experiment with a mixture of avatars with T1DM and real users/players, in which the players can compete with each other as well as the avatars, about being better regulated. This requires that the metabolism models and other physical models on which the avatars are based on should be as realistic as possible.

Conclusion

The results received along with the feedback from the user group indicate that the game has potential to be a useful tool for children and adolescents to learn about diabetes. We believe that this will be important to improve self-management

for children and adolescents with T1DM who live quite far from each other. Especially for adolescents, T1DM can be stigmatising. If their friends don't understand why they have to measure blood glucose level and inject insulin, it is sometimes socially difficult to do so. But a shared game experience may make it easier.

Further implementation and testing is of course needed to assure that the learning goals can be met, and that is also how we plan to continue going forward, until both we and our user base is sufficiently satisfied with the game as a tool for learning to manage diabetes.

Acknowledgments

References

[1] Chomutare T, Johansen S-G, Årsand E, and Hartvigsen G. Serious Game Co-design for Children with Type 1 Diabetes. *Studies in Health Technology and Informatics*, 2016;226:83-6. PMID: 27350472

[2] Makhlysheva A: A mobile phone-based serious game for children with Type 1 diabetes. Master's thesis in Telemedicine and e-Health. June 2013. University of Tromsø. (2013)

[3] Makhlysheva A, Årsand E, Varmedal R, Leknessund A, and Hartvigsen G. Use of Patient-Recorded Data in a Smartphone-based Game for Children with Diabetes. *Diabetes Technology & Therapeutics*. 2014, 16(S1): A-116 - A-117

[4] Rønningen IC: Exploring In-Game Rewards in the Diaquarium: A Serious Game for Children with Type 1 Diabetes Mellitus. Master's thesis in Computer Science. December 2016. University of Tromsø – The Arctic University of Norway. (2016)

[5] Rønningen IC, Årsand E, and Hartvigsen G. Exploring In-Game Reward Mechanisms in Diaquarium – A Serious Game for Children with Type 1 Diabetes. *Lecture Notes in Bioinformatics (LNBI) (Subseries of Lecture Notes in Computer Science)*. 2018, Vol. 10814, pp. 443-455. (ISSN 0302-9743)

[6] Årsand E, Makhlysheva A, Bradway M, Chomutare T, Johansen S-G, Blixgård H, and Hartvigsen G. Serious Gaming in Diabetes: Combining Apps and Gaming Principles in a Holistic Diabetes Environment. *Diabetes Technology & Therapeutics*. February 2016, 18(S1): A-86.

[7] Makhlysheva A, Årsand E, and Hartvigsen G. Review of Serious Games for People with Diabetes. Book chapter in: Novak, D., Tulu, B., Brendryen, H. (Eds.), Handbook of Research on Holistic Perspectives in Gamification for Clinical Practice. pp. 412-447). Hershey, PA: IGI Global, 2016. (DOI: 10.4018/978-1-4666-9522-1.ch019) (ISBN13: 9781466695221)

[8] Nobel Web AB. The Diabetic Dog Game. Available: http://nobelprize.org/educational_games/medicine/insulin/index.html 2010 (Cited: 18.05.18)

[9] Medtronic. Carb Counting with Lenny. Available: http://www.lenny-diabetes.com/carb-counting-with-lenny.html (2011) (Cited: 18.05.18)

[10] Nobelpriskampen: http://nobelpriskampen.se/2009/diabetes_insulin/ 2009. (Cited: 18.05.18)

[11] Johansen S-G, Makhlysheva A, Årsand E, Bradway M, and Hartvigsen G. Designing motivational and educational diabetes video games involving children as a creative resource. *Diabetes Technology & Therapeutics*. February 2016, 18(S1): A-92.

Address for correspondence

svein.gunnar.johansen@gmail.com

Use of Welfare Technology to Increase Employment of Individuals with Intellectual Disabilities

Sofie Wass[a], Carl Erik Moe[b], Elin Thygesen[c], Silje Haugland[d]

[a]*Department of Health and Sport Sciences, University of Agder, Grimstad, Norway*
[b]*Department of Information Systems, University of Agder, Kristiansand, Norway*
[c]*Department of Health and Nursing Science, University of Agder, Kristiansand, Norway*
[d]*Department of Psycosocial Health, University of Agder, Grimstad, Norway*

Abstract

Welfare technology can be applied to increase the involvement and independence of individuals with disabilities. While it is mainly applied for elderly, there are also initiatives for persons with intellectual disabilities, for different purposes. This group is currently marginalized in the labour market and there is a need to increase the support for employment. In this study, we provide an overview of previous literature reviews on intellectual disability and employment. Based on these findings, we discuss in which areas welfare technology could support employment of individuals with intellectual disabilities. The results show that employer attitudes, job coaches and support programs are important for employment. Drawing on prioritised areas within welfare technology, we recommend to study how technology can be supportive within these areas, focusing on social inclusion in working life, a structured working life and public service delivery.

Keywords:

Intellectual Disability, Employment, Technology.

Introduction

Welfare technology is seen as an important concept and innovation policy in the Scandinavian countries [1]. With an increasing need for welfare services, and with fewer people to provide those services, technology is viewed as an important step in managing that challenge. Welfare technology can be applied to maintain or increase involvement and/or independence of individuals with disabilities [2, 3]. It encompasses services for clients, healthcare professionals, relatives, industries and the society [2] and is seen as a heterogeneous group of technologies ranging from communication support, assistive technology, disease management, technology for everyday tasks, entertainment and social support [4, 5]. In Norway, welfare technology is often defined as *"...technological assistance that contributes to increased security, social participation, mobility and physical and cultural activity, and strengthens the individual's ability to manage himself in everyday life despite illness and social,* psychological or physical impairment. Welfare technology can also serve as support for next-of-kin and otherwise help improve accessibility, resource utilization and quality of service"* [3].

Welfare technology is mainly applied for elderly living at home, for instance as safety and fall alarms and different kind of sensors implemented in the home environment. Other examples include technology that provides medication reminders and the use of tablets and mobile phones to reduce social isolation and to increase efficiency [2, 3, 5]. However, there are also initiatives for applying welfare technology for persons with intellectual disabilities, both in Norway [3, 6] and in other Scandinavian countries [2]. These initiatives include sensors and alarms [2] but also technologies for localisation, communication, structure and time management and information exchange between different actors [7].

Today, a majority of Norwegian individuals with intellectual disabilities either have placements at day-centers (48%) or in workplaces provided by social care services (41%). In addition, almost all individuals receive social support at the age of 26 [8] and compared to other OECD countries there is a high rate of incapacity-related support [9]. A Norwegian report shows that individuals with disabilities are marginalized in both the traditional labor market and in segregated workplaces within the state labor market initiative [8]. A similar situation is also the case for other Scandinavian countries [10, 11]. This is a challenge as an active working life is described as one of the foundations for inclusion in society. Apart from earning livelihood, it has a positive impact on establishing a social network and identity, increasing self-esteem, providing structure and increasing health and well-being [10, 12-14]. Hence, we argue that there is a need to increase the support for employment of individuals with intellectual disabilities.

The work market is changing, it is becoming more unstable and complex, and also asks for flexibility of workers – and this makes it important to understand how individuals with intellectual disabilities can be supported in the work market [15]. The aim of this paper is to present an overview of previous literature reviews on intellectual disability and em-

ployment and to identify areas that are of importance for obtaining and maintaining employment. Based on these findings, we aim to discuss in which areas welfare technology could support the employment of individuals with intellectual disabilities.

Methods

The databases of Academic Search Complete, MedLine, PsycINFO, CINAHL and EMBASE were scanned for reviews focusing on intellectual disability and employment. The following keywords were used: intellectual disability AND employment AND review, and these were searched for in the abstract, with no restriction for included years. In total, 123 articles and book chapters were obtained. We included both systematic and scoping reviews, that investigated barriers and enablers for employment of individuals with intellectual disabilities. In addition, one review article was retrieved based on back tracking references, hence; we started with a total of 124 articles and book chapters.

After removing duplicates (n=76) and book chapters (n=4), the abstracts of 44 articles were read to determine an inclusion or not. This resulted in the exclusion of 39 articles due to a focus on the situation of individuals with disabilities in specific countries (n=5), specific approaches or interventions (n=8), cost analysis (n=4), quality of life or social inclusion (n=7), not providing a review of existing literature or lacking a description of the search strategy (n=11), other focus (n=4) or not being accessible (n=3). In total, 5 review articles were included for further analysis (Figure 1).
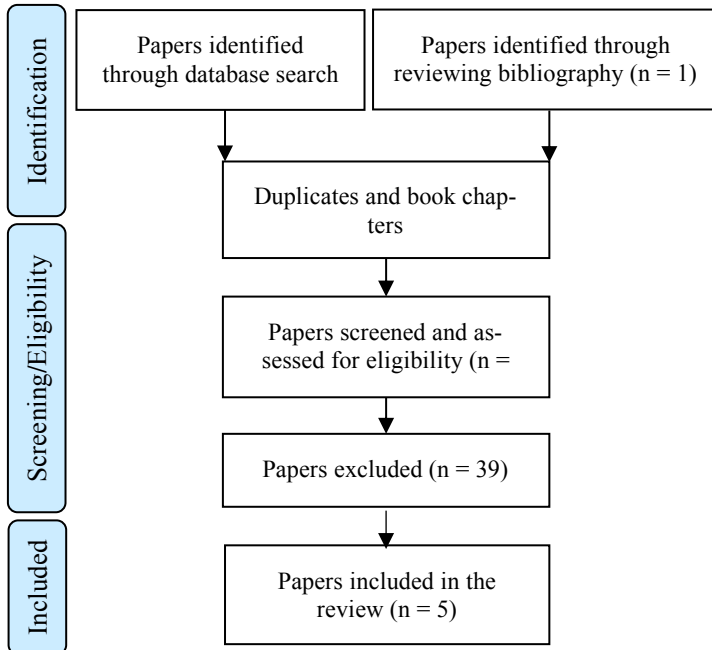
Figure - Flow diagram of the search process.

As a second step, the articles were read in detail. The cited studies, included in the reviews, that reported barriers and enablers for employment were classified in four main themes; the workplace context, the individual context, the societal context and the use of technology or techniques. The cited studies were assessed to determine the outcome of the investigated enablers/barriers. For example, in the review by Cheng, Oakman [16], a study by McInnes et al. [2010] reported that "*following job coaching, participants are three times more likely to be employed*". Hence, a positive impact was coded for the area individual context – job coaches. Studies that stated that employment of individuals with disabilities was low, but with no further explanation were not included in the analysis. Table 1 shows an overview of how the barriers and enablers were classified into different areas (a detailed overview can be obtained from the lead author).

Table 1: Areas that influence the employment of individuals with intellectual disabilities.

Area Study	[17]	[16]	[18]	[19]	[20]
Total number of studies	27	20	50	28	55
Workplace context = 35 articles					
Co-workers' support	10	4	4	-	6
Employer attitudes	4	-	-	7	-
Individual context = 51 articles					
Job training & Job search assistance	5	6	14	1	3
Job matching	1	1	2	3	7
Job coaches	3	4	1	-	-
Societal context = 8 articles					
Welfare benefits	1	-	-	3	4
Welfare technology & techniques = 18 articles					
Technology	-	3	9	-	-
Techniques	3	2	1	-	-
Studies not focusing on factors for employment	0	0	19	14	35

Results

The workplace context

The workplace context included issues connected to support from co-workers and opinions of employers regarding individuals with intellectual disabilities. The largest number of studies relating to the workplace context was included in the review by Ellenkamp, Brouwers [17] (n=14). While it was found that support from co-workers was important for integration and interaction in the workplace (n=10), it seems unclear how it affects the possibilities of obtaining or main-

taining employment. Only two studies found support for maintaining work. The results in the review by Cheng, Oakman [16] were varied, one of the studies cited found that support from co-workers can increase job placement rates while the other three studies found no support for such rates. Similar findings were indicated in the review by Hedley, Uljarevic [18] in which one study identified that a supportive workplace fosters the success of individuals with autism spectrum disorder. In two other studies, employees with intellectual disabilities stressed the importance of support and understanding in the workplace but the effect on the quality of the employment was unclear. Six studies in the review by Lövgren, Markström [20] suggested that peer support improves the opportunities to obtain and keep an employment but with no ascertained evidence. Two of the studies did stress that simply providing support is not enough, there is also need for some form of education, for instance mentoring, and for financial support. This was explored in another study. The study showed that while education in mentoring for staff members who work with job training for individuals with intellectual disabilities improved the feedback provided by staff members, the behavior of the individuals with disabilities remained unchanged [18].

Employers' attitudes towards employing individuals with disabilities were identified as an important enabler in two of the reviews [17, 19]. Two studies found that safety, productivity, attendance, availability of supportive services, no behavioral problems and punctuality are of importance to employers [17]. In addition, employers with previous experiences of staff with intellectual disability are more likely to employ individuals with disabilities compared to those without previous experiences [17]. The review by Cavanagh, Bartram [19] focused on human resource management and identified seven studies which reported that employer attitudes are a barrier towards employment of individuals with intellectual disability.

The individual context

A large number of articles focused on support for the individual. This included job training, job search assistance, job matching and job coaches. The review by Hedley, Uljarevic [18] found 14 studies which explored the importance of employment support programs and services, such as job search assistance and on-the-job training. The impact of such support was reported to have a positive effect on obtaining employment in all fourteen studies and one of the studies reported positive influence on increased working hours and wages. Four out of five studies in the review by Ellenkamp, Brouwers [17] found that job training was an important enabler for obtaining employment. The review by Cheng, Oakman [16] included six articles that focused on different kind of job support in which job search assistance and on-the-job training were found to be helpful in gaining and retaining an open employment. A small case study found support for the combination of off-the-job training and on-the-job training for individuals with autism for quicker skill uptake and gaining experience of a work context. While it made the participants work ready it is unclear if it had any effect on employment [19]. Alternatives to early placement

for on-the-job training were presented in one review (n=3) [20]. This included short periods at different workplaces for individuals with autism in order to foster an understanding of work. However; the connection to obtaining or maintaining employment was not reported.

The importance of matching the interests of the individual with the employment was mentioned in all the reviews [16-20]. One study found positive support for the use of person-centered planning to determine employment preferences [16] and another study found that matching interests and abilities increased the possibility to maintain employment [17]. Two studies recommended that matches should be more strategic and focus on occupations where individuals with intellectual disabilities are well-represented [20]. Ten additional studies recommended matching of interests and work, still it seems unclear how it affects the possibilities of obtaining or maintaining employment.

Four studies in the review by Cheng, Oakman [16] focused on job coaches. Three of them showed that coaches have a positive effect on obtaining and retaining an employment. This was also indicated by four other studies which found that job coaches were important for employers' decision to hire individuals with intellectual disabilities and for maintaining employment [17, 18]. However, one study showed that a decrease of job coach support can have positive outcomes on work productivity for individuals with severe intellectual disability [16].

The societal context

One study related to welfare benefits showed that subsidies for employers as well as individually placed persons increased the salary of individuals with intellectual disabilities [17]. The role of welfare programs was also mentioned in the review by Cavanagh, in which three studies reported on the need for increased support from such programs. A call for more coordinated work around welfare programs was mentioned in four other studies [20]. However, the negative impact on employment due to current programs was not described.

Welfare technology and techniques

Concerning completion of work tasks, three of the reviews included studies that focused on the use of instructional approaches (n=6) and welfare technologies (n=12) [16-18]. Different kind of non-technological approaches included training for work tasks with the use of specific words, career development tasks [17] and checklists [16] which all showed to increase the work performance of individuals. One study discussed an assessment instrument [17] and two articles explored behavior techniques including incidental teaching, discrete-trail teaching and social stories [16, 18]. These results seemed promising but due to limited research, the effects on employment outcomes had not been identified.

The use of welfare technology to assist individuals with disabilities included teaching of tasks through e.g. video instructions and audio coaching [16] which both showed to improve work performance. Hedley, Uljarevic [18] also reported on positive outcomes of implementing a personal digital assis-

tant which increased working time and using virtual jobs which improved interviewing skills. The use of an iPad at work was also shown to increase independence, confidence, time management and organizational skills. However, the effects of video self-modelling to learn work tasks [16] and the impact of other applications could not be ascertained [18]. In total, half of the studies did not report a positive impact on improved work performance while the other half did.

Discussion

The reviews included in our study focused on barriers and enablers for employment of individuals with intellectual disability. The reviews showed limited support for positive outcomes on employment for areas such as co-workers' support and job matching. While these areas were stressed as important in several studies, few studies reported on actual impact on either obtaining or maintaining employment. On the other hand, the support from co-workers did however increase the integration of individuals with intellectual disabilities and their interaction with colleagues. A few studies discussed welfare benefits and criticized current initiatives but lacked support for its' negative impact on employment.

The reviews and the cited studies focusing on employer attitudes showed that their attitudes towards individuals with disabilities are important for obtaining and maintaining an employment. While there were few studies focusing on job coaches, seven out of eight studies supported the positive influence of job coaches. In addition, the majority of the studies on support programs (22/27) reported positive evidence related to employment. The studies of applications of different approaches and technologies to support employment of individuals with intellectual disabilities, reported mixed results, showing either positive effects on work performance or the need of more research. To summarize, the following areas seem to be important for the ability to obtain and maintain an employment: (1) employer attitudes, (2) job coaches and (3) support programs.

It appears that the employment of individuals with intellectual disabilities depends on several initiatives and that there is a need to work within several areas. In view of the current discussion on the potential of welfare technology [21] it is of vital importance to explore if and how welfare technology can be applied within these areas. In Norway, an influential white paper argued to focus on three main technology areas: safety alarms, technology that increases social inclusion and technology that supports an active and structured everyday life [3]. For individuals with intellectual disabilities, the last two areas are of interest due to the current exclusion from working life [8, 9, 15] and due to the individuals' disability, limiting their actions that are performed within a societal environment [22, 23]. Drawing on these recommendations and the results in our study, we recommend to explore the potential of welfare technologies in the following areas:

- Technologies that support social inclusion in working life, targeting employer attitudes, activities provided by job coaches and activities performed in support programs.

- Technologies that support a structured working life, targeting activities provided by job coaches and activities within support programs.

Apart from the three main areas, there is a need to invest in technologies that support the delivery of public services and the exchange of information between all involved actors in the welfare system [3]. This is similar to what is described as back-office technologies for public e-services, i.e. the parts of the service process which is not visible to citizens but connected to the technology in the supplying organization [24]. With the wide range of actors involved in the use of public e-services and its complexity, it is important that information can be communicated and shared without disruption [25]. We therefore propose to study technologies that support the internal components of public services that are offered to individuals with intellectual disabilities. This includes technologies used by other users, for instance those who are employed in social services, schools and healthcare services, who are in contact with individuals with intellectual disabilities.

Limitations

This study included English language reviews published in peer-reviewed outlets, which neglects findings presented in white papers and in other languages. Further, only a small number of reviews were identified, however; we believe that the included reviews cover many studies and that our study presents a synthesis of the literature.

Conclusions

Our synthesis of the reviews and the included articles presents a number of areas that are of importance for employment of individuals with intellectual disabilities. Employer attitudes, job coaches, support programs and, to some extent, technology appear to play an important role for the possibility to obtain and maintain employment. With this background, it should be further explored how welfare technologies can be applied within these areas, focusing on social inclusion, a structured working life and public service delivery.

References

[1] Brynn R. Universal Design and Welfare Technology. Studies in health technology and informatics 2016: 229: 335-44.

[2] Socialstyrelsen. E-hälsa och välfärdsteknik i kommunerna 2017.

[3] Norges offentlige utredninger. Innovasjon i omsorg. 2011.

[4] Hofmann B. Ethical Challenges with Welfare Technology: A Review of the Literature. Sci Eng Ethics. 2013: 19 (2): 389-406.

[5] Stowe S, Harding S. Telecare, telehealth and telemedicine. Eur Geriatr Med. 2010: 1 (3): 193-7.

[6] Knarvik U, Rotvold G-H, Bjørvig S, Bakkevoll P-A. Kunnskapsoppsummering: Velferdsteknologi. Tromsø: Nasjonalt senter for e-helseforskning, 2017; pp. 1-43.

[7] Trondsen MV, Knarvik U. Velferdsteknologi for barn og unge med funksjonsnedsettelser. Tromsø: Nasjonalt senter for e-helseforskning, 2017; pp. 1-34.

[8] Wendelborg C, Kittelsaa AM, Wik SE. Overgang skole arbeidsliv for elever med utviklingshemming. Trondheim, Norge: NTNU Samfunnsforskning AS, 2017.

[9] OECD. Investing in Youth: Norway Paris: OECD Publishing, 2018.

[10] Tideman M, Lövgren V, Szönyi K, Bergkvist C, Haghjo A. Intellektuell funktionsnedsättning och arbete. Stockholm, Sweden: Forte, 2017.

[11] Børne- og Socialministeriet. Modtagere af kontanthjælp med handicap. In: Afdeling for Analyse og Datastrateg, editor, 2018; pp. 1-12.

[12] Kober R, Eggleton IRC. The effect of different types of employment on quality of life. J Intell Disabil Res. 2005: 49 (10): 756-60.

[13] Law M, Steinwender S, Leclair L. Occupation, Health and Well-Being. Can J Occup Ther. 1998: 65 (2): 81-91.

[14] Beyer S, Brown T, Akandi R, Rapley M. A Comparison of Quality of Life Outcomes for People with Intellectual Disabilities in Supported Employment, Day Services and Employment Enterprises. J Appl Res Intellect. 2010: 23 (3): 290-5.

[15] Santilli S, Nota L, Ginevra MC, Soresi S. Career adaptability, hope and life satisfaction in workers with intellectual disability. J Vocat Behav. 2014: 85 (1): 67-74.

[16] Cheng C, Oakman J, Bigby C, Fossey E, Cavanagh J, Meacham H, et al. What constitutes effective support in obtaining and maintaining employment for individuals with intellectual disability? A scoping review. J Intellect Dev Dis. 2017: 1-11.

[17] Ellenkamp JJ, Brouwers EP, Embregts PJ, Joosen MC, van Weeghel J. Work Environment-Related Factors in Obtaining and Maintaining Work in a Competitive Employment Setting for Employees with Intellectual Disabilities: A Systematic Review. J Occup Rehabil. 2016: 26 (1): 56-69.

[18] Hedley D, Uljarevic M, Cameron L, Halder S, Richdale A, Dissanayake C. Employment programmes and interventions targeting adults with autism spectrum disorder: A systematic review of the literature. Autism. 2017: 21 (8): 929-41.

[19] Cavanagh J, Bartram T, Meacham H, Bigby C, Oakman J, Fossey E. Supporting workers with disabilities: a scoping review of the role of human resource management in contemporary organisations. Asia Pac J Hum Resou. 2017: 55 (1): 6-43.

[20] Lövgren V, Markström U, Sauer L. Towards Employment: What Research Says About Support-to-Work in Relation to Psychiatric and Intellectual Disabilities. J Soc Work Disabil Rehab. 2017: 16 (1): 14-37.

[21] Nordens välfärdscenter. Nordisk tankesmedja om välfärdsteknologi 2018 [Available from: https://nordicwelfare.org/projekt/nordisk-tankesmedja-om-valfardsteknologi/.

[22] ICD-10. Den internasjonale statistiske klassifikasjonen av sykdommer og beslektede helseproblemer Diktoratet for e-helse, 2018.

[23] Association AP. Diagnostic and statistical manual of mental disorders (DSM-5®): American Psychiatric Pub, 2013.

[24] Lindgren I, Jansson G. Electronic services in the public sector: A conceptual framework. Gov Inform Q. 2013: 30 (2): 163-72.

[25] Axelsson K, Melin U, Lindgren I. Public e-services for agency efficiency and citizen benefit — Findings from a stakeholder centered analysis. Gov Inform Q. 2013: 30 (1): 10-22.

Address for correspondence

Sofie Wass, sofie.wass@uia.no

Predicting Cost-effectiveness of Telehealthcare to Patients with COPD: A Feasibility Study Based on Data from the TeleCare North Cluster-randomized Trial

Flemming Witt Udsen, Ole Hejlesen

Department of Health Science & Technology, Aalborg University, Aalborg, Denmark

Abstract

International results have recently questioned the value of providing telehealthcare to all COPD patients. Results from the Danish TeleCare North trial nuanced the debate by concluding that telehealthcare would most likely only be cost-effective for patients in a subgroup of severe COPD. Machine-learning methods have been suggested as a strategy to target telehealthcare even better than clinical subgroups. Data from the TeleCare North trial was used to fit classification models in order to explore this feasibility. Three models were applied: a simple decision tree, logistic regression and a linear support vector machine. Results indicate that classification methods can be used to predict patient-level cost-effectiveness with a relatively high precision. With these methods, it is feasible to target telehealthcare even better in order to maximize survival and health-related quality of life while not overusing scarce health resources as argued by health economists and clinical advocates of rational medicine.

Keywords:

Pattern Recognition, Automated/classification; Cost-Benefit Analysis, Telemedicine; Pulmonary Disease, Chronic Obstructive, Denmark.

Introduction

Chronic obstructive pulmonary disease (COPD) is a prevalent disease [1,2]. The World Health Organization projects that 64 million people have COPD worldwide, attributing to more than 3 million deaths annually [3]. COPD is a major health problem in the European Union with a member state COPD prevalence between 4-10% [4]. The Danish Lung Association estimates that 430,000 Danish citizens have COPD [5] with a scientific study setting the prevalence to 9% [6]. According to an analysis of the disease burden in Denmark, around 3,300 Danish citizens die due to COPD each year (6% of total number of deaths) [7] and 2,500 die each year of causes related to COPD [8].

Systematic reviews and a recent large-scale randomized trial have shown some effect of telehealthcare on health outcomes [9–11]. However, due to fiscal pressure in healthcare sectors, effective interventions might still not receive public funding, if there is not a reasonable relationship between health outcomes and associated costs [12]. Although several pilot trials have suggested that telehealthcare could be cost-effective or cost-saving for COPD patients [13–17], recent larger-scale studies have found it hard to replicate similar results, suggesting that it might not be cost-effective to scale up telehealthcare solutions to all COPD patients assuming that "one-solution fit all" [18–21].

However, there is little knowledge on the implications of patient heterogeneity, i.e. that some patient groups might for some reason be more or less cost-effective than others. Recent results from the Danish TeleCare North trial is one of the first studies to indicate that, although telehealthcare is not cost-effective for all COPD patients (21), it might – on average - be highly cost-effective for patients in the severe COPD subgroup [22].

But at the same time, there is a growing demand for targeting telehealthcare even better than by clinical subgroups in order to identify the contexts and characteristics of ideal individual patients that should receive telehealthcare, i.e. enabling assessment that answers under which circumstances telehealthcare would be most likely to be cost-effective [23–25]. Machine-learning methods used to predict outcomes for individual patients eligible for telehealthcare has recently been suggested as one way forward [26].

The aim of this project is to investigate if it is possible to use predictive algorithms to help stratify telehealthcare for COPD patients in a way that maximizes the patient-level cost-effectiveness ratio.

Materials and Methods

Data

The study design, data collection procedure [27] as well as the results of TeleCare North has been been published elsewhere [21,22]. In addition to usual care, patients in the intervention group received a set of telehealthcare equipment, disease-specific education and were monitored by a municipality-based healthcare team. The control group received usual care. 1,225 patients were included with 578 patients

receiving telehealthcare and 647 usual care. The duration of the study was 12 months.

For the purposes of this study, the variables collected as part of the study is of particular interest and can be divided into the cost-effectiveness ratio to be predicted and a set of features used to predict this outcome.

Outcome

The primary outcome is the total costs per quality-adjusted life year (QALY) gained at the individual level, which represent an individualized cost-effectiveness ratio.

A QALY is a composite measure of survival and health-related quality-of-life (HRQoL) [28]. Information on mortality was taken from the Danish Register of Causes of Death [29]. HRQoL stemmed from generic EQ-5D-3L questionnaires distributed to patients at baseline and 12-months follow-up. The EQ-5D scores HRQoL on a scale from 0-1, where "1" indicates perfect HRQoL and "0" meaning dead. Negative scores are possible indicating health states considered worse than death [30]. QALYs were calculated by linear interpolation of EQ5D-3L scores with Danish societal weights [31,32].

Total costs included intervention costs, healthcare costs (hospital-, medicine-, and primary sector costs), and social sector costs (cost to practical help and care at home, home-based nursing care, and rehabilitation). Within-trial healthcare costs were all collected from national registers by applying patients' unique social security number. Hospital contacts were collected from the Danish National Patient Register [33]; contacts between patients and the primary care sector from the National Health Insurance Service Register [34]; and medication use was taken from The Danish Register of Medicinal Product Statistics [35]. Social sector costs was estimated from electronic care systems in each of the 10 included municipalities. Intervention costs were assessed from prices paid during the TeleCare North implementation and included costs of hardware, installation, maintenance and support as well as training costs, monitoring costs, and project management costs. All costs were reported in 2014 prices and were obtained in Danish kroner (DKK) and thereafter exchanged to € using the average 2014 exchange rate (1€=7.4547 DKK).

Features

Different baseline features were collected as part of the trial. All features originated from a combination of questionnaires, physical measurements conducted by general practitioners and from national registers. Three baseline HRQoL summary scores (baseline EQ-5D, PCS and MCS scores) was available from two generic questionnaires distributed to all patients: the EQ-5D-3L [30] and the SF-36 [36]. Six physiological parameters was measured by the patients' general practitioner and consisted of systolic- and diastolic blood pressure, pulse, body mass index (BMI), spirometry measures (percentage of expected forced expiratory volume in one second (FEV1 (%)), percentage of expected forced vital capacity (FVC (%)). Six socio-demographics (marital status, highest education, duration of COPD, job status, number of persons

in household and smoking status) and the presence of five comorbidities (diabetes, musculoskeletal disease, cancer, mental illness and heart disease) was ascertained from questionnaires that were filled out by patients. Age and gender was identified from the patients' social security number and their residing municipality was collected from the Danish Civil Registration System [37]. Nine features associated with the historical activity for the included patients was incorporated to account for organizational differences and variation in visitation practices. Therefore, the number and duration of hospital admissions and the number of outpatient visits one year prior to randomization was included as were the costs associated with these contacts. Costs due to medicine, primary care and practical help and care at home, home-based nursing care, and rehabilitation was also collected for the one year leading up to the evaluation period. All features was collected from the same national registers as described above.

In total, the dataset consist of 32 features. In addition, six predictors were derived from other features: In the original cost-effectiveness analysis, three characteristics were important in distinguishing between subgroups that were more cost-effective than others [22].The cost-effectiveness of telehealthcare depended on COPD severity group that can be calculated from FEV1(%) [38]. Furthermore, existing social sector costs was a driver of cost-effectiveness; as were baseline cost (accumulated costs one year prior to evaluation). Both variables were calculated from the individual baseline cost-categories that were collected as part of the trial. Three other features were generated that indicate whether a patient suffer from hypertension, multimorbidities or tachycardia and they were calculated from blood pressure, individual comorbidities and pulse, respectively. These were included because of their clinical meaningfulness.

Missing data

Of the 1,225 patients originally included in the TeleCare North trial, complete data for both total costs (i.e. all cost categories), baseline EQ-5D score and EQ-5D score at follow-up were available for 728 patients (59%; 302 in telehealthcare group; 426 in control group). Missing data for the EQ-5D summary score were present for 8% of the participants at baseline (48 in the telehealthcare group; 53 in the control group). Due to non-response or to incomplete registration of EQ-5D questionnaire items, 27% had missing data on the EQ-5D summary score at follow-up (199 in the telehealthcare group; 133 in the control group). Two municipalities where unable to extract rehabilitation costs (79 in the telehealthcare group; 73 in the control group).

Only the 302 participants with complete outcome data in the telehealthcare group were included in this analysis.

Missing data in any features from those 302 participants were replaced by multiple imputation. Missing feature values were assumed missing at random (MAR) at replaced with the *mi impute chained* command in STATA12.1 and 30 complete datasets were created an averaged to one value per missing. Continuous variables were imputed by predictive mean

matching and categorical variables by multinomial logistic or logistic regression. Imputation models included predictors for the outcomes at both time points and predictors for missing observations in the individual variables. The imputation models were estimated separately by treatment group and included the clustering variable and the measures of disease status, presence of comorbidities and sociodemographic variables described previously.

The characteristics of the included patients are presented in Table 1.

Table 1- Baseline characteristics of the sample (n=302)

Age (years)	69.57 (8.87)
Men (%)	50.99 (n=154)
Marital status	
Married/in a relationship	66.56 (n=201)
Single	17.55 (n=53)
Widow/widower	15.89 (n=48)
Smoking status	
Non-smokers	65.89 (n=199)
Smokers	34.11 (n=103)
Highest education	
Elementary school	46.03 (n=139)
Secondary school	4.30 (n=13)
Vocational education	34.77 (n=105)
Short tertiary school (2-3 years)	7.95 (n=24)
Bachelor or equivalent (3-5 years)	6.95 (n=21)
Master or equivalent (>5 years)	0.00 (n=0)
Job status	
Full-time	3.64 (n=11)
Part-time	7.28 (n=22)
None	89.07 (n=269)
Blood pressure, diastolic	76.64 (10.51)
Blood pressure, systolic	130.62 (16.97)
Pulse	79.12 (13.31)
BMI	26.22 (5.01)
Duration of COPD (years)	7.97 (6.10)
FEV1 (%)	47.57 (16.29)
FVC (%)	70.85 (16.94)
Comorbidities	
Diabetes	10.69 (n=32)
Heart disease	32.12 (n=97)
Mental health problem	3.31 (n=10)
Musculoskeletal disease	25.83 (n=78)
Cancer	5.63 (n=17)
Baseline total cost (€)	4,863.37 (7,874.23)
Baseline HRQoL	
EQ-5D score	0.727 (0.20)
PCS (SF-36 physical score)	38.30 (8.68)

MCS (SF-36 mental score)	48.98 (10.97)

Data are mean (SD) or proportion (number of patients)
COPD: Chronic obstructive pulmonary disease; FEV1(%): forced expiratory volume in one second of predicted normal; FVC(%): forced vital capacity

Model and features

To investigate some basic methods of machine-learning, a relative naïve approach was chosen for this study. The prediction of cost-effectiveness was addressed as a simple pattern classification problem to test different methods.

First, all 38 features regardless of multicollinearity were used in subsequent analyses.

Second, the individualized cost-effectiveness ratio was categorized into two groups (Class 1: Total cost/QALY ≤ €5,000; Class 2: Total cost/QALY>€5,000). This was done due to simplicity and because the cost-effectiveness ratio was highly skewed even after normalization. The boundary was exploratively chosen based on the distribution of the individualized cost-effectiveness ratio.

Third, MATLAB Release 2017a was used to train and evaluate all 38 predictors using a five-fold cross-validation. A simple decision tree, logistic regression and linear support vector machines were fitted to the data.

Evaluation

All models were evaluated based on accuracy and the area under the receiver operating characteristics curve (AUC). Accuracy is the proportion of correct predictions made for the two cost-effectiveness categories among the total number of observations. The AUC curve quantifies the overall ability of the classification model to discriminate between observations that have a total cost/QALY ≤ €5,000 or >€5,000 across classification thresholds. A model with no better precision than chance has an AUC of 0.5.

Results

Table 2 presents model performance for each of the three models that were fitted. The linear SVM performed best with an accuracy of 79.1% and an area under the curve of 0.89. The classification model was therefore able to classify 79% of the observations into the correct cost-effectiveness categories with a relatively high ability to distinguish which observations should be predicted to have an individualized cost-effectiveness larger or smaller than €5,000 across classification levels.

Moreover, all three models had relatively high performance (accuracy between 76.7-79.1% and area under the curves between 0.81-0.89).

Table 2- Performance evaluation of the included models

	Accuracy (%)	AUC
Simple tree	76.7	0.81

Logistic regression	77.4	0.84
Linear SVM	79.1	0.89

AUC: Area under the curve; SVM: Support vector machine

Figure 1 and Figure 2 presents the receiver operation characteristics of both classes (Class 1: Total cost/QALY≤ €5,000; Class 2: Total cost/QALY>€5,000) for the best performing classification model (linear SVM).
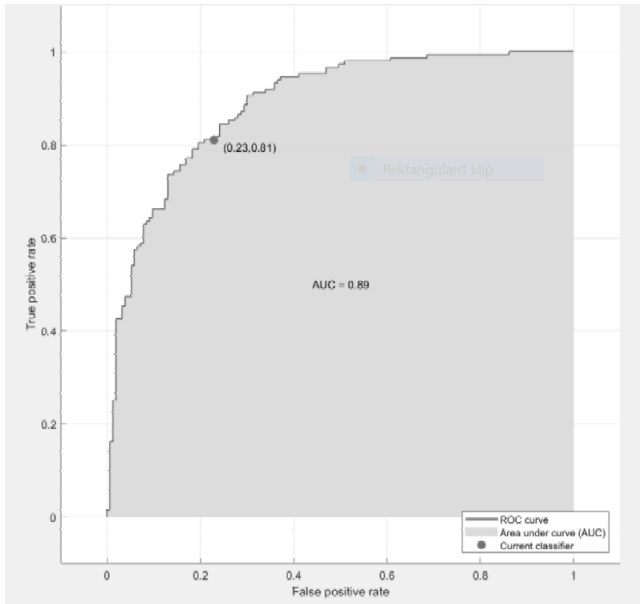
Figure 1- Receiver operation characteristics based on a linear support vector machine (Class 1)
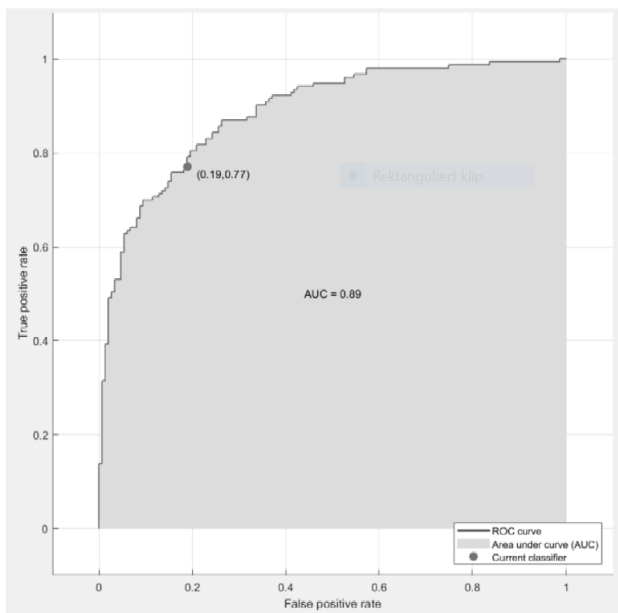
Figure 2- Receiver operation characteristics based on a linear support vector machine (Class 2)

Discussion

The best performing model was the linear SVM, but all three compared models had relatively high accuracy (from 76.7-79.1) and large area under the ROC-curves (from 0.81-0.89). This seems to suggest that machine-learning methods can be used to predict the individualized cost-effectiveness of tele-health to COPD patients with a relatively high precision.

Strength/weaknesses

This study applied data from a large clinical trial that rigorously collected patient-level data on both health outcomes, socio-demographics and resource patterns across all relevant stakeholders involved in managing COPD patients. All cost data are based on unambiguous register-data complete eliminating any self-report biases.

A limitation is that no feature selection was conducted prior to model training. Using 38 features to predict a two-category outcome resulting in a relatively high precision might be due to overfitting of the data, which means that the models accurately classifies the training data, but would perform poorly on future data. A five-fold cross validation was used to counter some of the problems of overfitting the data, but in general, much fewer features could make the same classification. From the original cost-effectiveness analysis, it is known that especially baseline HRQoL and total baseline costs (the sum of costs accumulated from all cost-categories 12 months prior to evaluation start) had a major impact on the cost-effectiveness estimate [21]. For this reason baseline EQ-5D is required as a variable in any models estimating cost-effectiveness in health economic research [39].

Another limitation is that no techniques for dealing with outliers were applied. It could rightfully be stated that very large outliers are attributed to patients having other serious disease alongside COPD. However, models did account for major comorbidities and baseline total costs. And health economists will insist on asking what threshold levels constitutes an outlier given smooth cost-effectiveness distributions and argue that outliers are crucial for estimating impact on healthcare budgets, which is why mean values instead of medians are always reported [40,41].

Finally, large proportions of missing data can be a particular problem in machine-learning. In clinical trials, good practice implies handling missing data with multiple imputation [42,43]; but this procedure could artificially boost the performance of prediction models, since multiple imputation entails using other variables to predict missing values. This study therefore sought to minimize the missing data challenge by focusing on the telehealthcare group with complete cost-effectiveness ratio. But these participants still had missing data on especially socio-demographics and physical measurements used to predict cost-effectiveness.

Comparison with other research

While some studies have applied machine-learning techniques in medicine to predict outcomes related to health eco-

nomic evaluation, i.e. mortality in heart disease [44], HRQoL in COPD [45] and even the cost of treatment in liver disease [46]; no other studies have been found that sought directly to predict the cost-effectiveness ratio of health technologies.

The study published by Stausholm and colleagues [44] come closest to the design of this study. It applied data from the TeleCare North trial to predict increasing or decreasing HRQoL and healthcare sector costs for 553 COPD patients receiving both telehealthcare and usual care. The design entailed application of four different logistic regression models with 39 features, which were evaluated by accuracy and root mean square error (RMSE). Accuracy of models ranged from 61-65% for models predicting HRQoL and 74-75% for models predicting healthcare sector costs, while the RMSE was 5.265 scores for HRQoL and $5430 for healthcare costs. The study concluded that predictive analytics could be used to stratify COPD patients for telehealthcare. This current study has a slightly higher accuracy (76.7-79.1%) with roughly the same number of features. However, this study applied a different HRQoL summary score (the EQ-5D as opposed to SF-36 in Stausholm) and accounted for a much broader cost-perspective. Especially costs in municipalities (due to daily practical help and home nursing care) led to larger proportion of COPD patients to have very large costs, that were in part accounted for by including a feature for having resource use in municipalities prior to data collection.

Implications

Based on the cost-effectiveness results from the TeleCare North trial, the Danish Government decided to fund telehealthcare to the subgroup of patients with severe COPD throughout Denmark [47]. Although, this subgroup was likely to be cost-effective on average, there is both a scientific and practical need to select COPD patients from an even more limited set of potential candidates for telehealthcare [26]. This would result in a quicker and less expensive implementation of telehealthcare that would facilitate better health outcomes while also reducing overall health- and social care costs.

Machine-learning methods already has an established role in healthcare and biomedicine [48] in areas such as disease discovery, e.g. [49] and diagnosis [50]. Results from this study seems to indicate that it would be possible to predict individual cost-effectiveness of telehealthcare for future patients. This would enable more personalized COPD management, i.e. a COPD management strategy that are based on individual data paired with a predictive toolbox that evaluates which combinations of individuals and technologies, that are eligible for receiving public funding.

Future research

More studies on the type of models used to predict cost-effectiveness should be conducted. The individualized cost-effectiveness ratio is a very left-skewed ratio-scaled variable ranging, in principle, from $-\infty$ to $+\infty$, but with mostly positive values. This study applied a pattern recognition strategy by dividing cost-effectiveness into two categories, but models that are used to predict the continuous cost-effectiveness ra-

tio or even more cost-effectiveness categories could have been applied instead.

New predictive models might also need to be developed and incorporated into existing software that can account for the clustered nature of health- and cost data: due to variation in practice, it is plausible that mortality, health-related quality of life and costs are more similar within e.g. geographical areas or the patients' organizational affiliation than across these areas. Treating this information merely as fixed variables in modelling outcomes can lead to biased model coefficients and –uncertainty [51,52].

Competing interest

None declared.

References

[1] Halbert RJ, Natoli JL, Gano A, Badamgarav E, Buist AS, Mannino DM. Global burden of COPD: systematic review and meta-analysis. Eur Respir J [Internet]. 2006;28(3):523–32. Available from: http://www.ncbi.nlm.nih.gov/pubmed/16611654

[2] Mannino DM, Buist AS. Global burden of COPD: risk factors, prevalence, and future trends. Lancet [Internet]. 2007;370(9589):765–73. Available from: http://www.ncbi.nlm.nih.gov/pubmed/17765526

[3] WHO. Chronic obstructive pulmonary disease (COPD) [Internet]. Fact sheet 315. 2012. Available from: http://www.who.int/mediacentre/factsheets/fs315/en/index.html

[4] European COPD Coalition. Prevalence of COPD in EU [Internet]. 2014. Available from: http://www.copdcoalition.eu/about-copd/prevalence

[5] Danish Lung Association. COPD in Denmark [Internet]. 2012. Available from: www.lunge.dk

[6] Hansen JG, Pedersen L, Overvad K, Omland Ø, Jensen HK, Sørensen HT. The Prevalence of chronic obstructive pulmonary disease among Danes aged 45-84 years: population-based study. COPD [Internet]. 2008;5(6):347–52. Available from: http://www.ncbi.nlm.nih.gov/pubmed/19353348

[7] Flachs EM, Eriksen L, Koch MB. The disease burden in Denmark. 2015.

[8] National Institute of Public Health. The Public Health Report, Denmark. 2007.

[9] McLean S, Nurmatov U, Liu JLY, Pagliari C, Car J, Sheikh A. Telehealthcare for chronic obstructive pulmonary disease: Cochrane Review and meta-analysis. Br J Gen Pract [Internet]. 2012;62(604):e739–49.

Available from: http://www.embase.com/search/results?subaction=viewr ecord&from=export&id=L365948923

[10] Steventon A., Bardsley M, Billings J, Dixon J, Doll H, Hirani S, et al. Effect of telehealth on use of secondary care and mortality: findings from the Whole System Demonstrator cluster randomised trial. Bmj [Internet]. 2012 Jun 21;344(jun21 3):e3874–e3874. Available from: http://www.bmj.com/cgi/doi/10.1136/bmj.e3874

[11] Wootton R. Twenty years of telemedicine in chronic disease management--an evidence synthesis. J Telemed Telecare [Internet]. 2012;18(4):211–20. Available from: http://www.pubmedcentral.nih.gov/articlerender.fcgi?art id=3366107&tool=pmcentrez&rendertype=abstract

[12] Drummond M, SculpherJ. M, Glaxton K, Stoddart L. G, Torrance W. G. Methods for the Economic Evaluation of Health Care Programmes. 4th edition. Oxford University Press; 2015.

[13] Haesum LKE, Soerensen N, Dinesen B, Nielsen C, Grann O, Hejlesen O, et al. Cost-utility analysis of a telerehabilitation program: A case study of COPD patients. Telemed e-Health [Internet]. 2012;18(9):688–92. Available from: http://www.scopus.com/inward/record.url?eid=2-s2.0-84869057985&partnerID=40&md5=87200bfb992018c29 819b0f75c0d60cf

[14] de Toledo P, Jiménez S, del Pozo F, Roca J, Alonso A, Hernandez C. Telemedicine experience for chronic care in COPD. IEEE Trans Inf Technol Biomed [Internet]. 2006 Jul;10(3):567–73. Available from: http://www.ncbi.nlm.nih.gov/pubmed/16871726

[15] Vitacca M, Assoni G, Pizzocaro P, Guerra A, Marchina L, Scalvini S, et al. A pilot study of nurse-led, home monitoring for patients with chronic respiratory failure and with mechanical ventilation assistance. J Telemed Telecare [Internet]. 2006;12(7):337–42. Available from: http://www.scopus.com/inward/record.url?eid=2-s2.0-39049188481&partnerID=40&md5=cb0d21f36d8f7c730 d9c5d2ff6b08297

[16] Koff PB, Jones RH, Cashman JM, Voelkel NF, Vandivier RW. Proactive integrated care improves quality of life in patients with COPD (Provisional abstract). Eur Respir J [Internet]. 2009 May [cited 2013 Sep 6];33(5):1031–8. Available from: http://onlinelibrary.wiley.com/o/cochrane/cleed/articles/ NHSEED-22009102080/frame.html

[17] Pare G, Poba-Nzaou P, Sicotte C, Beaupre A, Lefrancois E, Nault D, et al. Comparing the costs of home telemonitoring and usual care of chronic obstructive pulmonary disease patients: A randomized controlled trial. Eur Res Telemed [Internet]. 2013;2(2):35–47. Available from: http://www.embase.com/search/results?subaction=viewr ecord&from=export&id=L52635611

[18] Henderson C, Knapp M, Fernández J, Beecham J, Hirani S, Cartwright M, et al. Cost effectiveness of telehealth for patients with long term conditions (Whole Systems Demonstrator telehealth questionnaire study): nested economic evaluation in a pragmatic, cluster randomised controlled trial. BMJ. 2013;346.

[19] Stoddart A, van der Pol M, Pinnock H, Hanley J, McCloughan L, Todd A, et al. Telemonitoring for chronic obstructive pulmonary disease: a cost and cost-utility analysis of a randomised controlled trial. J Telemed Telecare [Internet]. 2015;21(2):108–18. Available from: http://jtt.sagepub.com/lookup/doi/10.1177/1357633X145 66574

[20] McDowell JE, McClean S, FitzGibbon F, Tate S. A randomised clinical trial of the effectiveness of home-based health care with telemonitoring in patients with COPD. J Telemed Telecare [Internet]. 2015;21(2):80–7. Available from: http://jtt.sagepub.com.ez.statsbiblioteket.dk:2048/conten t/21/2/80.full

[21] Udsen F, Lilholt P, Hejlesen O, Ehlers L. Cost-effectiveness of telehealthcare to patients with chronic obstructive pulmonary disease: Results from the Danish "TeleCare North" cluster-randomized trial. BMJ Open. 2017;7(5).

[22] Udsen FW, Lilholt PH, Hejlesen OK, Ehlers L. Subgroup analysis of telehealthcare to patients with chronic obstructive pulmonary disease: The case of the Danish "TeleCare North" cluster-randomized trial. Clin Outcomes Res. 2017;9:391–401.

[23] Dixon P, Hollinghurst S, Edwards L, Thoms C, Foster A, Davies B, et al. Cost-effectiveness of telehealth for patients with depression: evidence from the Healthlines randomised controlled trial. BJPsych Open [Internet]. 2016;2:262–269. Available from: http://bmjopen.bmj.com/lookup/doi/10.1136/bmjopen-2016-012355

[24] Dixon P, Hollinghurst S, Edwards L, Thomas C, Gaunt D, Foster A, et al. Cost-effectiveness of telehealth for patients with raised cardiovascular disease risk: evidence from the Healthlines randomised controlled trial. BMJ Open [Internet]. 2016;6(9):e012355. Available from: http://bmjopen.bmj.com/lookup/doi/10.1136/bmjopen-2016-012355

[25] Udsen F. Health Economic Evaluation of Telehealthcare: Can we include "why" and "under what circumstances" telehealthcare is cost-effective in health economic evaluation? [Internet]. Aalborg: Aalborg University Press; 2016. Available from: http://aauforlag.dk/Shop/phd-udgivelse/phd-by-flemming-witt-udsen.aspx

[26] Vitacca M, Montini A, Comini L. How will telemedicine change clinical practice in chronic obstructive pulmonary disease? Ther Adv Respir Dis [Internet]. 2018;12:175346581875477. Available from: http://journals.sagepub.com/doi/10.1177/1753465818754 778

[27] Udsen F, Lilholt P, Hejlesen O, Ehlers L. Effectiveness and cost-effectiveness of telehealthcare for chronic obstructive pulmonary disease: study protocol for the Danish "TeleCare North" pragmatic cluster-randomized trial. Trials. 2014;15(178).

[28] Brazier J, Ratcliffe J, Salomon JA, Tsuchiya A. Measuring and Valuing Health Benefits for Economic Valuation. 1st ed. New York: Oxford University Press;

2009. 344 p.

[29] The State Serum Institute. The Danish Register of Causes of Death. 2014.

[30] Reenen M van, Oppe M. EQ-5D-3L User Guide. EuroQol Res Found. 2015;(April):22.

[31] Sørensen J, Davidsen M, Gudex C, Pedersen KM, Brønnum-Hansen H. Danish EQ-5D population norms. Scand J Public Health. 2009;37:467–74.

[32] Wittrup-Jensen K, Lauridsen J, Gudex C, Pedersen K. Generation of a Danish TTO value set for EQ-5D health states. Scand J Public Health. 2009;37:459–66.

[33] Lynge E, Sandegaard JL, Rebolj M. The Danish National Patient Register. Scand J Public Health. 2011;39(7 Suppl):30–3.

[34] The State Serum Institute. National Health Insurance Service Register [Internet]. 2014. Available from: http://www.ssi.dk/Sundhedsdataogit/Registre og kliniske databaser/De nationale sundhedsregistre/Sundhedsokonomi finansiering/Sygesikringsregister.aspx

[35] Wallach Kildemoes H, Toft Sorensen H, Hallas J. The Danish National Prescription Registry. Scand J Public Health. 2011;39(7 Suppl):38–41.

[36] Ware J, Sherbourne C. The MOS 36-item short-form health survey (SF-36). I. Conceptual framework and item selection. Med Care. 1992;30(6):473–83.

[37] Pedersen CB, Gøtzsche H, Møller JO, Mortensen PB. The Danish Civil Registration System. A cohort of eight million persons. Dan Med Bull. 2006;53(4):441–9.

[38] Guide P, Copd TO. Global Initiative for Chronic Obstructive Lung A Guide for Health Care Professionals Global Initiative for Chronic Obstructive Disease. Glob Initiat chronic Obstr lung Dis. 2010;22(4):1–30.

[39] Manca A, Hawkins N, Sculpher MJ. Estimating mean QALYs in trial-based cost-effectiveness analysis: The importance of controlling for baseline utility. Health Econ. 2005;14(5):487–96.

[40] Drummond M, Sculpher M, Claxton K, Stoddart G, Torrance G. Methods for the Economic Evaluation of Health Care Programmes. Fourth edi. Oxford University Press; 2015.

[41] Glick H, Doshi J, Sonnad S, Polsky D. Economic Evaluation in Clinical Trials. Oxford University Press; 2007.

[42] Faria R, Gomes M, Epstein D, White IR. A guide to handling missing data in cost-effectiveness analysis conducted within randomised controlled trials. Pharmacoeconomics [Internet]. 2014;32(12):1157–70. Available from: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=4244574&tool=pmcentrez&rendertype=abstract

[43] Sterne JAC, White IR, Carlin JB, Spratt M, Royston P, Kenward MG, et al. Multiple imputation for missing data in epidemiological and clinical research : potential and pitfalls. 2009;339(July):157–60.

[44] Motwani M, Dey D, Berman DS, Germano G, Achenbach S, Al-Mallah MH, et al. Machine learning for prediction of all-cause mortality in patients with suspected coronary artery disease: A 5-year multicentre prospective registry analysis. Eur Heart J. 2017;38(7):500–7.

[45] Stausholm MN, Egmose A, Dahl SC, Lilholt PH, Cichosz SL, Hejlesen OK. Stratification of telehealthcare for patients with chronic obstructive pulmonary disease using a predictive algorithm as decision support: A pilot study. J Telemed Telecare [Internet]. 2017;23(3):410–5. Available from: http://journals.sagepub.com/doi/10.1177/1357633X16644129

[46] Bertocco L, Haddad DP. Machine Learning in the Prediction of Costs Open for Liver Transplantation. 2017;25:335–43.

[47] Finansministeriet. Aftaler om den kommunale og regionale økonomi for 2016 [Internet]. 2015. 90 p. Available from: http://www.fm.dk/publikationer/2007/aftaler-om-den-kommunale-og-regionale-oekonomi-for-2008/download/~/media/Files/Publikationer/2008/Download/Aftale_om_den_kommunale_og_regionale_okonomi_2008_endelig_version.ashx

[48] Yoo I, Alafaireet P, Marinov M, Pena-Hernandez K, Gopidi R, Chang JF, et al. Data mining in healthcare and biomedicine: A survey of the literature. J Med Syst. 2012;36(4):2431–48.

[49] Vijeta S, Ajai K, Lakshmi P, Ganesh K, Anuradha L. Malaria Outbreak Prediction Model Using Machine Learning. Int J Adv Res Comput Eng Technol [Internet]. 2015;4(12):4415–9. Available from: http://ijarcet.org/wp-content/uploads/IJARCET-VOL-4-ISSUE-12-4415-4419.pdf

[50] Gunčar G, Kukar M, Notar M, Brvar M, Černelč P, Notar M, et al. An application of machine learning to haematological diagnosis. Sci Rep. 2018;8(1):1–12.

[51] Gomes M, Ng ES-W, Grieve R, Nixon R, Carpenter J, Thompson SG. Developing appropriate methods for cost-effectiveness analysis of cluster randomized trials. Med Decis Making [Internet]. 2012 [cited 2014 Oct 16];32(2):350–61.

[52] Rabe-Hesketh S, Skrondal A. Multilevel and Longitudinal Modeling Using Stata. Volume I: Continouos Responses. Third edit. Texas: Stata Press; 2012.

Address for correspondence

Flemming Witt Udsen, Department of Health Science & Technology, Aalborg University, Fredrik Bajers Vej 7C, DK-9220 Aalborg Ø, Denmark.
E-mail: fwu@hst.aau.dk

Designing a Dashboard to Visualize Patient Information

Janus Waidtløv Gustafson, Camilla Holt Jones, Louise Pape-Haugaard

Department of Health Science and Technology, Aalborg University, Aalborg, Denmark

Abstract

Patient handovers from prehospital to emergency departments (ED) can be complex; involving critical patient care under the influence of high stress levels and carries the potential for loss of important contextual information. An intuitive and easy access to this information, can advantageously help ensure patient safety. The aim of this study was to design a dashboard prototype that visualizes data for the clinicians in emergency departments in acute care. Through observations at an emergency department, a dashboard design was conducted and evaluated. Six clinicians from two different emergency departments and three peers with experience from healthcare, were used as evaluators of the dashboard, by a cognitive walkthrough using a case-based simulation. This study found that the evaluators perceived the visualization of the patient information at the dashboard, as easily manageable and sufficient. Visualizing important contextual patient information at a dashboard, can be rewarding to the clinicians in the ED.

Keywords:

Data visualization, Emergency Medical Services, Emergency Department, Clinical Decision-Making, Dashboard, Design.

Introduction

Prehospital information involves contextual information about the incident location, patient status, vital signs etc. [1,2]. Information can both support the clinicians at an emergency department and ensure patient safety, if properly used [1,3]. The complex and dynamic situations in patient handovers from prehospital to emergency departments put high demands for fast decision making about further procedures and treatments [1]. Decision making processes are challenged by high levels of stress. High levels of stress are a psychological factor that may impact the ability to transform information to actions [1]. Hence, decision making relies on easily accessible contextual information, which should be visualized intuitively "at a glance" [1,4,5].

Visual simplistic projecting of data facilitates information processing to be faster [1,6]. The visualization of data can be efficiently displayed by a dashboard design, the design promotes interrelations between data. Today, information about the patient can be given in writing or orally to an emergency coordinator, or directly presented in a complex information system [3].

The aim of this study was to design a dashboard prototype that simplistically visualizes and prioritizes data for clinicians in emergency departments in acute care. The design was further evaluated to consolidate the possible effects of the design.

Materials and Methods

The study consisted of two phases:
1. Designing the electronic dashboard.
2. Evaluation of the clinical dashboard

Phase 1: Designing the clinical dashboard

To investigate the needs for a simplistically design of the dashboard prototype and to form the design process insights into everyday practices were obtained by qualitative observations. Analysis of the workflow provided an intuitive understanding of possible solutions, as well as knowledge of how the clinicians currently uses prehospital information. In addition, a detailed description of the physical conditions framed the workflow.

The observation was conducted in an emergency department (ED) by two observers, over the course of one day and was conducted as an observational study with the possibility of elaborate questions to the clinicians. The observation and dialogues were used as possibilities for optimizing data placement in the dashboard design.

The dashboard design was conducted on the basis of *the good design principles* as described by Wiklund et al. [7] in comparison with the results from the field studies and outcomes from other studies [2,4–6,8,9]. Based on this, the results were 20 requirements (see table 1) which were set in the user interface and content which became the basis for the design.

The prototype was designed using interactive mockups in the prototype tool Justinmind Prototypes (Ver 8.3.1).

Phase 2: Evaluation of the electronic dashboard

The prototype was evaluated by a cognitive walkthrough, which involved descriptions of the workflow utilizing case-based simulation [7]. During the walkthrough an in-depth description of the prototype was presented, thereby

increasing the motivation and rational thinking of the evaluators [10].

The evaluation consisted of three stages:

1. Case presentation. The evaluators were asked to indicate the relevant information.

2. Dashboard was displayed. The evaluators were asked to select which data was relevant, as well as what alternatives could be desired.

3. Opinion regarding; access to data, design and whether the dashboard appeared to be easy or difficult to use.

The procedure was the same for all evaluators, and each research team member had the same role in each evaluation. The dashboard was presented in the same manner to all evaluators. Firstly, the data panels were presented, secondly navigational options and at last a header containing data which was not relevant in acute treatments of emergency patients.

Evaluators

Nine evaluators were selected, to evaluate the dashboard. Three evaluators from ED A (EDA), in which the field studies where held. Three evaluators were also selected from the Department of Health Science and Technology, Aalborg University, Aalborg, Denmark (peers), as well as three evaluators from ED B (EDB) which is located in another region. A total of nine evaluators were selected, all with a bachelor's degree within health science and at least two years of health care experience. The evaluators from the two emergency departments were expected to provide important information, concerning the how design was experienced, and whether the data displayed on the dashboard harmonized with the workflow and the information needed in emergency medicine.

Results

The trauma room had installed in it a big screen which was visible to all attending personnel. This is ideal for showing the latest data but was currently only used for showing pictures from the accident scene.

Twenty requirements to the dashboard were constructed, as illustrated in table 1. The requirements are based on the observation findings, outcomes from other studies and the principles as described by by Wiklund et al [7]. The dialogue with the clinician concluded that notes from the emergency respondent, vital signs and pictures from the accident scene was of highest importance.

The twenty requirements were utilized in the design of the dashboard design.

The dashboard consists of 3 graphical parts:

1. Data panels consisting of six panels shown in figure 1.

2. A navigational cluster in which the user can shift between interfaces.

3. A header in the upper part of the user interface which contains multiple kinds of data.

Table 1- Requirements for the electronic dashboard design

Requirements for the electronic dashboard design			
1	Must fit multiple monitor sizes.	11	Data must be shown close together on the user interface in panels.
2	Alerts for incoming critical patient's must remain visible.	12	There must be consistency in the chosen design i.e. same placement of controls with same features throughout the prototype.
3	Important data should be visible on dashboard.	13	Icons and titles must have relational meaning to each other.
4	Panels containing new data should be placed at the top of screen.	14	Navigation should be through a few buttons placed in a cluster.
5	If data exceeds screen size only vertical scrolling should be promoted.	15	The dashboard must show as much data as possible, so navigation is minimized.
6	Login periods should be prolonged.	16	Patient identification must be placed in a header that is visible at all times.
7	The critical and important data should be visible on dashboard.	17	Titles should be meaningful to the user.
8	Critical data must be emphasized.	18	Navigational buttons should be both icons and titles.
9	Continuously synchronization of data with timestamp.	19	Overview should be simplified with few panels containing data.
10	User interface must be compatible with multiple platforms.	20	The graphical user interface should be aesthetically pleasing.

Walkthrough of the data panels

The arrangement of the panels is sorted by its content and relation to each other (see figure 1, in Danish). The six prioritized panels contain:

1. **Patient** - The panel contains patient data such as allergies, previous diseases, diagnosis and identification-related data

2. **Assessment and treatment** - The panel contains Airway, Breathing, Circulation, Disability, Exposure (ABCDE), as well as treatments performed by the prehospital unit.

3. **Vital signs & observations** - The panel contains observations and measured vital signs that are critical for the treatment of the patient. These are divided into columns so that they are clearly separated from each other.

4. **Pictures** - The panel contains images from the scene of injury and of the patient.

5. **Notes** - The panel contains all notes taken by the pre-hospital unit.

6. **Injuries** - The panel contains a diagram of the patient's injuries.

The panels facilitate visualization of the user interface and sought-after data is quickly and intuitively accessible. Most

recent data are prioritized to the top of each panel, ensuring uniform data presentation across all panels. Each value/ observation has its own column to provide an overview. This minimizes the possibilities for misinterpretation of data. Clinical decision support is solved by gathering data into few panels where relational data is gathered in the same panel. i.e. ABCDE is shown with patient treatment.

Figure 1- The figure shows the six data panels used to visualize data

To meet the requirements concerning information density, the panels have clear borders which gives the user a reading direction and keeps the design from looking cluttered. Information are therefore easily distinguished from each other. The amount of data generated in the prototype has not filled out the panels, therefore scrolling has not been required. Attempts have been made to create a high degree of internal consistency by placing features in the same place across all interfaces and keeping the design as uniform as possible. This applies more specifically to the features "expansion" and "minimize" of panels.

Evaluation findings

Evaluators' perception of the dashboard was noted for each individual during the cognitive walkthrough. The evaluators had a predominantly positive image of the dashboard.

- **Patient panel** - Few comments on this panel. A single evaluator from EDB commented that it was good to know if a patient's identity was confirmed by themselves, so that the clinician was prepared for the actual condition of the patient.

- **Assessment and Treatment** – The evaluators had some remarks, and as a peer expressed it, it is difficult to comprehend so much text in one panel. An evaluator from EDB made the point that the dashboard design puts pressure at the prehospital units, filling the ABCDE scores.

- **Vital signs and observations panel** - Multiple evaluators appreciated the presented values in columns as designed in the prototype, so that development could be followed. Two evaluators indicated that they would like to concurrently see the values on a curve. When asked whether they preferred the values shown numerically or placed in a curve, all stated they would prefer the current numerical representation. One evaluator pointed out that, critical values could be marked red to draw the clinician's attention.

- **Pictures panel** – Evaluators from EDA were surprisingly not excited about pictures from the scene of injury. This was especially interesting as they currently used the pictures. Evaluators from EDB understood the idea and thought well of it. The peers group having no domain experience, needed an explanation of why the images were of importance. A single evalua-

tor from this group thought the picture was big and her gaze were drawn there.

- **Notes panel** - Notes from prehospital units are currently used by EDA but not by EDB. The benefits of notations from the prehospital was acknowledged by both departments.

- **Injuries panel** – All evaluators were pleased with the injury picture. Especially by the possibility to expand the injury picture, so a description of the damage could be seen. One evaluator made the point, that labels could be difficult to differentiate, if a multi trauma with many types of injuries were shown.

- **General Overview** - All evaluators felt that the dashboard provided a nice overview of relevant clinical data. One of the peers felt like there was too much information in the dashboard but had no experience within the domain of emergency medicine, and therefore lacks the insight into what information is considered being important in trauma situations.

- **Other comments** – Evaluators from EDA appreciated that only data fields containing information were shown in the dashboard, this eliminates noise in the overall impression of the dashboard. Evaluators from EDB suggested easier access to the information system. Their suggestions included a longer login period and a joint user, so they did not have to spend time on logging into the system. A single evaluator from EDB also wanted the screen in the trauma room to be split into two, so the notes (as they received by telephone from the prehospital units) could be displayed at the same time as the dashboard.

Discussion

The use of a dashboard improves the ability of health professionals to effectively find information and improves the sharing of information [5]. The results of this evaluation support this. The use of data panels containing relevant contextual data has proven to be beneficial for the clinicians. Most evaluators found the data easy to find.

The results show that dashboard designs can provide a good overview of complex data and how this data should be distributed across a dashboard used in emergency departments. This is supported by other studies (9,11,12)

The study design was conducted in a user centered approach. The field study provided knowledge about users, their workflows, and their challenges with the current use of prehospital patient information. Based on this, a prototype dashboard was prepared to meet the needs of the clinicians. The design process was iterative, alternating between data collection from field studying and involvement of relevant outcomes from other studies

The study involved multiple iterations concerning relevance and inspiration, but only one design iteration, which can be seen as a limitation. This design iteration involved relevant

users from an emergency department, the involvement of relevant users, who have the possibility to provide feedback, is positive for the use of a design, as well as users acceptance [13]. The involvement of users from the emergency departments is seen as an advantage in this study, due to their greater knowledge in the domain of emergency medicine and the various work procedures. It is essential that continuous involvement of end users right from the beginning of the design process and preferably with the same facilitators [14].

The field study identified that the trauma room settings were ideal for the use of a dashboard design and a dialogue with the attending nurse led to the conclusion that notes from the prehospital respondent, vital signs and pictures from the accident scene were of highest importance and thus became vital parts of the data panels.

- **Vital signs and observations** - It can be seen as important that data is presented, in both numeric values and with visual illustrations [1]. In this study, it was not possible to present vital signs visually by a curve on the main screen, but as numerical values arranged in columns. Data should not require a cognitive transformation to be understood by the user, as data may otherwise be misunderstood in stressful situations [1].

The cognitive walkthrough of the prototype found that the clinicians preferred numerical data, with the possibility of accessing a visualization as needed. The evaluators found the data as sufficient with vital signs arranged in columns and arranged chronologically with the latest measurement at the top. Data can be presented in multiple ways, as an evaluator suggested could the data fields alert the user if a value exceeds a limit. However, additional studies would be required, as more abnormal values would cause more red markings, allowing abstraction from other panels, which are also important in the acute situations [7].

- **Notes** -Throughout the cognitive walkthroughs of the dashboard, the use of the contextual data in the dashboard was explained by the same approach, but evaluators' perception of affordance did not seem to be the same. Affordance is the way in which objects are perceived by the evaluator's own habits and assumptions. Obtaining affordance can be met by an intuitive design, where the user is not in doubt about the application [1,15,16]. Affordance was not achieved especially by one evaluator. The evaluator did not perceive that the dashboard interaction was with prehospital units under the "notes" panel but would like to have it simultaneously display the emergency department's own notes from their electronic health records. A reason for this could be that the dashboard title *notes* did not appear to have an intuitively meaning for this evaluator.

- **Pictures** - While evaluating the prototype, one evaluator found that the pictures from the accident site appeared to be a focal point. This is consistent with the cognitive mindset of user interface design, where

the eye first sees the graphics, then highlighted text and finally the body text [17,18]. Despite the relevance not found in the literature and similar studies, the cognitive walkthrough showed the importance of having pictures in the data panels. However, two factors that may be related to accident pictures; estimated accident rate and damage mechanism [2]. Pictures from the accident site describe the kinetics of the accident, which serves the purpose of the two factors. However, a description of the two factors imposes higher demands on prehospital units than taking pictures, and time management is likely to be necessary.

In conclusion, this study sought to design and evaluate a simplistically prototype dashboard that visualizes data for the clinicians in emergency departments in acute situations like traumas through an analysis of clinicians' requirements and needs in combination with *good design principles* by Wiklund [7]. The overall positive evaluation shows that contextual information displayed in a dashboard is appreciated by the clinicians and is useful in critical situations.

Limitations

The cognitive walkthrough was chosen as an evaluation technique because the aim of the study was to design a dashboard visualizing the relevant contextual patient data, in correlation with the workflow at an ED. The cognitive walkthrough showed to be ideal. However, the evaluations were conducted under artificial conditions, away from the environment in which the dashboard is expected to be used. It was not possible to test the dashboard under trauma treatment conditions i.e. in trauma room. Alternatively, the evaluation could have been conducted as a usability study, where two parallel, simulated and real-life workflows are performed, starting from the incoming trauma patient for treatment at the trauma room (19). This approach would seem more genuine for the clinicians and the dashboard could be tested under conditions that seem more realistic. At the same time, the current use of prehospital data and the prototype could be assessed in relation to each other.

References

[1] Steiner CM, Nussbaumer A, Neville K, Albert D. A Psychological Framework to Enable Effective Cognitive Processing in the Design of Emergency Management Information Systems. Electron J Inf Syst Eval. 2017;20(1):39–54.

[2] Carter AJE, Davis KA, Evans L V., Cone DC. Information Loss in Emergency Medical Services Handover of Trauma Patients. Prehospital Emerg Care. 2009;13(3):280–5.

[3] Siemsen IMD, Madsen MD, Pedersen LF, Michaelsen L, Pedersen AV, Østergaard D, et al. Factors that impact on the safety of patient handovers: An interview study. Scand J Public Health. 2012;40(5):439–48.

[4] Swartz JL, Cimino JJ, Fred MR, Green RA, Vawdrey DK. Designing a clinical dashboard to fill information gaps in the emergency department. AMIA Annu Symp Proc. 2014;2014:1098–104.

[5] Dowding D, Randell R, Gardner P, Fitzpatrick G, Dykes P, Favela J, et al. Dashboards for improving patient care: Review of the literature. Int J Med Inform. 2015;84(2):87–100.

[6] Franklin A, Gantela S, Shifarraw S, Johnson TR, Robinson DJ, King BR, et al. Dashboard visualizations: Supporting real-time throughput decision-making. J Biomed Inform. 2017 Jul;71:211–21.

[7] Wiklund ME, Limor JK, Ul H, Weinger MB. Technical Basis for User Interface Design of Health IT. 2015;15–996.

[8] Batley NJ, Osman HO, Kazzi AA, Musallam KM. Implementation of an emergency department computer system: Design features that users value. J Emerg Med. 2011;41(6):693–700.

[9] Rasmussen R. Electronic whiteboards in emergency medicine: A systematic review. Proc 2nd ACM SIGHIT Int 2012;483–92.

[10] Rubin J, Chisnell D. Handbook of usability testing : how to plan, design, and conduct effective tests. Wiley Pub; 2008. 348 p.

[11] Batley NJ, Osman HO, Kazzi AA, Musallam KM, Lee K, Jung SY, et al. A novel concept for integrating and delivering health information using a comprehensive digital dashboard: An analysis of healthcare professionals' intention to adopt a new system and the trend of its real usage. Chan, V and DAmbrogio, A and Zacharewicz, G and Mustafee, N, editor. Ann Emerg Med. 2017 Mar 2;32(1):98–108.

[12] Lee K, Jung SY, Hwang H, Yoo S, Baek HY, Baek RM, et al. A novel concept for integrating and delivering health information using a comprehensive digital dashboard: An analysis of healthcare professionals' intention to adopt a new system and the trend of its real usage. Int J Med Inform. 2017 Jan;97:98–108.

[13] Gagnon MP, Ghandour EK, Talla PK, Simonyan D, Godin G, Labrecque M, et al. Electronic health record acceptance by physicians: Testing an integrated theoretical model. J Biomed Inform. 2014;48:17–27.

[14] Sligo J, Gauld R, Roberts V, Villa L. A literature review for large-scale health information system project planning, implementation and evaluation. Int J Med Inform. 2017;97:86–97.

[15] Rosli DI. Cognitive Awareness Prototype Development

on User Interface Design. 2015;14(2):32–40.

[16] Kurosu M. Human-Computer Interaction. 2015. 320-331 p.

[17] Salman YB, Cheng H-I, Patterson PE. Icon and user interface design for emergency medical information systems: a case study. Int J Med Inform. 2012 Jan;81(1):29–35.

[18] Lin H, Hsieh YC, Wu FG. A study on the relationships between different presentation modes of graphical icons and users' attention. Comput Human Behav. 2016;63:218–28.

[19] Gao T, Massey T, Selavo L, Welsh M, Sarrafzadeh M. Participatory User Centered Design Techniques for a Large Scale Ad-Hoc Health Information System. 2007;43–8.

Address for correspondence

Janus Waidtløv Gustafson januswgustafson@gmail.com

Motivation in Self-monitoring Processes:
Evaluation of Ecological Momentary Storytelling

Katja Lund[a], Lisbeth Kappelgaard[b]

[a] *Department of Electronic Systems, Aalborg University, Aalborg, Denmark*
[b] *Department of Communication and Psychology, Aalborg University, Aalborg, Denmark*

Abstract

Patient-involved treatment, such as self-monitoring, is a central ambition in health care in Scandinavia. Norway, Sweden, England and the Netherlands have enacted legislation on the patient's right to involvement [1]. In Denmark, patient involvement is formulated as one of the ten national health goals [2]. Patient-governed treatment is politically articulated as a way to individual empowerment [3]. Nevertheless, doctors and nurses find that many patients express reluctance and a lack of motivation [4]. In line with the political discourse and with an ambition to uncover work-related stressors the authors developed the self-monitoring method Ecological Momentary Storytelling. The purpose of the article is to present test participants' articulated experiences with using the method. Through a grounded theory based analysis of follow-up dialogues with the participants, the findings emphasize how motivation is not solely anchored inside the individual as a personal desire to master a situation or be empowered. Matters such as complex life situations, inability to handle technology and problems understanding questions posed are to a large extend externally anchored and articulated as hindrances for motivation.

Keywords:

Occupational Stress, Ecological Momentary Assessments, Telemedicine, Sense of Coherence, Grounded Theory, Hearing Loss, Motivation, Methodological Study

Introduction

The use of self-monitoring is playing an increasing role as a solution to prevention and treatment in the Scandinavian healthcare system. Based on a broader understanding of health and illness, which has evolved through World Health Organization (WHO) strategies for health in recent decades, self-monitoring has a strong ideological link to the concept of empowerment [5]. Through the focus on empowerment, the use of self-monitoring is driven by the logic that patients should be experts on their own diseases and that this expertise leads to empowerment [6,7]. Self-monitoring is politically formulated as a way patients can take care of themselves and be released from time-consuming medical visits. The overall goal is for the patient to achieve a greater degree of autonomy and thus, be less dependent on hospital and health professionals [6].

In addition to benefitting the patient, self-monitoring is argued to comply with economic demands in society as successful patient involvement is aimed at qualifying the treatment, reducing professionals' workload and thus, reducing healthcare expenses [8]. Evidence for the effect of various self-monitoring methods is extensive in the literature, for example, in connection with the treatment of chronic disorders [9] and mental illnesses [10,11]. In much of this literature, researchers conclude that these new treatments have the potential to make the patient a "master of [his or her] own disease" [11–13].

Despite the asserted evidence-based effects, the empowering visions and economic achievements, many self-monitoring initiatives end up as pilot projects and do not become a robust part of daily practices [14]. One reason is that patients are reluctant and lack motivation [4]. In this sense, it seems that there is a discrepancy between the political discourse and practice [15].

A search of the literature on motivational factors and reluctance to practice eHealth, mHealth and self-monitoring shows that the field appears to be dominated by almost everything except a focus on long-term motivation—and especially spontaneous reluctance [16]. The majority of studies focus on eHealth in relation to clinical outcomes and less on patient engagement [17]. In literature dealing with patient engagement in connection with self-governed treatment, the terms used to describe patient engagement are diverse. Patient engagement, or patient activation, has become a generally accepted umbrella term, which positions patients in a central role in their own care [17]. Patient engagement considers patients as consumers involved in a specific sociocultural context as the term is derived from marketing literature [18], and patient engagement, driven by inner motivation, is increasingly considered a crucial factor in the quality of health care [19,20].

The purpose of this article is to evaluate Ecological Momentary Storytelling – a method for self-monitoring experiences of stress. Our ambition is to contribute to the field of engagement and motivation in self-monitoring processes by analyzing which factors the participants in the study articu-

lated as determinant to their positive and negative experiences with using the method. We therefore ask:

Through which norms and interpretative frameworks do the participants understand the monitoring and its output? How is this interpretation related to either staying motivated or developing reluctance?

Background

The self-monitoring method, Ecological Momentary Storytelling, was developed as a result of a collaboration between the authors, who shared an interest in occupational stress. One project had a special focus on communication and stress among hearing impaired people in the Danish work force, while the second project aimed at identifying which discourses on work-related stress were produced in the professional field of teachers.

In addition to having work-related stress as a common subject field, the authors had both a methodical ambition and a methodical frustration that coincided. Both projects originally set out in a qualitative method: We wanted to talk our way into the core of stressors - interact with the relevant actors and through reflective dialogues gain insight in experienced stress issues. Both authors aimed that the studies should give voice to those who have tried practice and possibly stressfulness on their own bodies. After having had several conversations with both teachers and employees with hearing loss, we shared the same thoughts and frustrations. We had gained insight into practice, but at the same time, we had an experience that many of the conversations were sensibility produced retrospectively. The stories outlined events that were often months or years old – told many times before. A story built on memory and 'backbone sense' can be fruitful and contain qualities - our memory enables us to actualize a forgotten knowledge and if it was not for this ability to forget and remember we would be left to the madness [21]. However, much research indicates that there are differences between the narratives produced from memory and narratives that reflect spontaneous here-and-now reactions to being in practice [22–25]. How could we deal with these differences? How did we capture both types of narratives? Another problem in relation to understand and capture stressors was a basic assumption that there may also be a silent knowledge of stress - for example bodily reactions that are not observable. The bodily signals can only be felt by the individual - it is a 'private language' [26] or a 'silent knowledge' [27], which is not always articualted. Therefore, it was crucial for us to develop a method that could provide insight into both spoken and silent knowledge [28]. These challenges became the starting point for the development of the self-monitoring method Ecological Momentary Storytelling.

In developing the method the authors shared an understanding of stress as a phenomenon with several, fundamentally inseparable, dimensions and the development of the stress tracking method was based on a holistic, interdisciplinary and bio-psycho-social stress concept [29]. In an attempt to reflect this understanding of stress, we developed a data-triangulation method, Ecological Momentary Storytelling, in which the theoretical inspiration was taken from the linking of EMA (ecological momentary assessment) [24] such as ESM (experience sampling method) and HRV (heart rate variability measssurements), medical sociology and humanistic psychology. The authors also adapted a salutogenetic approach to stress and coping [30], which in recent years in health practices in Denmark has been widely applied especially within the nursing area [31].

The Ecological Momentary Storytelling method consisted of three main pillars:

1. Reflective Dialogues: To make the data-logs accessible to the participants themselves we had to create a space for them where they could move from tacit to explicit knowledge. A start-up and follow-up dialogue were implemented as a part of the method. In our understanding of the dialogue concept, we relied on humanistic psychology. We were particularly inspired by Kristiansen and Bloch-Poulsen [32] who define dialogues as unpredictable, risky and exploratory conversations, where there is no predetermined truth and where creations are produced in the interpersonal contact. The goal was to open up new insights and opportunities together. Central to this dialogue understanding is that it is a special way of being present - inspired by Carl Rogers' three relational concepts: "Empathy, congruence and unconditional positive regard" [33].

2. Ambulatory monitoring: Drawing on existing research on ambulatory monitoring in relation to stress tracking [i.e. 33,34] we chose to log HRV-measurements to inform us on physical reactions and possible bodily experiences of stress. The HRV-data was used in a qualitative way where peaks and deviances in the data were related to the contexts and the ESM-data and reflected upon by the participants themselves in the reflective dialogues.

3. ESM (experience sampling method): The test persons logged their here-and-now experiences with ESM on a smartphone. This way psychological and also social aspects of a situation were logged and reflected upon in the dialogues.

The authors integrated the salutogenetic perspective into the content design by translating the three dimensions of the SOC (sense of coherence): 1) manageability, 2) comprehensibility and 3) meaningfulness [36] into questions on here-and-now experiences of inner balance, overview, and meaningfulness. This was done in order to evaluate the overall ability to handle stressors in the moment of logging. To develop the content of the ESM we used design methods based in participatory design thinking such as a type of cultural probing inspired by a Dutch study on hospital reality from a lying perspective [37] and material storytelling [38]. Through these activities, persons in the target group helped define central issues to address in the ESM. The final design included

logs on mood, energy- and noise-level, experiences with communication and number of people in the room (see Picture 1-3). To anchor each log to something that would support a fairly precise mental reproduction of the different situations during the week and thus lower the risk of memory bias when reflecting on the context of the separate logs in the dialogue session, the option for taking a photo and / or recording 10 seconds of sound was present. The method was developed and user-tested during 2012 – 2013 [12].
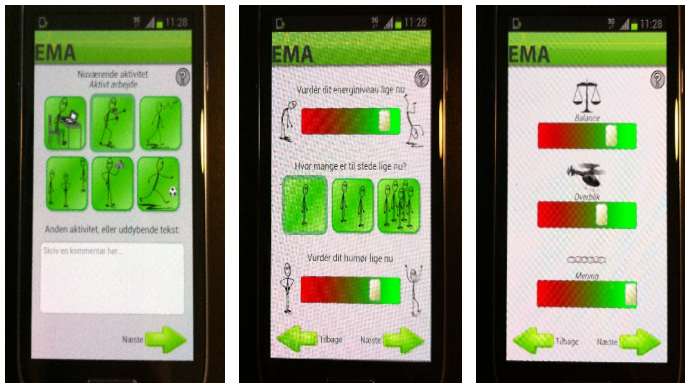
Picture 1-3: The images show screens of the application, where activities and experiences can be logged throughout the day. Picture 1 (from the left) shows the screen where you can register the activity you are engaged in, as well as add text. Picture 2 (in the middle) shows the screen where you can log the experience of the energy level, the number of people in the room and the mood. Picture 3 (right) shows the screen where you can log SOC by recording an assessment of whether you feel balanced, have an overview and feel what you are doing is meaningful.

Participants

Contact to the 48 persons with hearing loss who volunteered in joining the project was established through the National Hearing Association in Denmark. The association informed their members about the opportunity to take part in the study, and interested members then contacted the project by e-mail. Eight participants were picked from the group. Participant criteria was that there should be some degree of hearing loss present and that the person should be engaged in work if not for full hours then at least some days during the week. This standard was important because a focal point of the study was to examine the effect caused by a hearing loss in work situations as a part of daily life. The eight hearing-impaired participants were between ages 43 and 64. Two participants withdrew from the study after only a couple of days due to technical challenges, and therefore the participant age for the remaining six ended up being 50-64 years. The results from this study therefore indicate contexts of importance particularly among this group but may be significant to the entire group of people with hearing disabilities in the working age. Three men and three women with hearing loss were represented in the study. The six participants were engaged in the study one week each during 2013 and 2014 and the data ma-

terial exceeds 2000 hours of HRV-measurements, experience-loggings and follow-up dialogues.

Method

The data, which is analyzed in this article, is based on the dialogues following a week of data collection with the Ecological Momentary Storytelling method [39]. In the follow-up dialogues, the participants reflected on daily activities based on the HRV and ESM data. To reach an understanding of the multiple contexts that affected momentary experiences throughout a day, we decided to code the transcribed dialogues using grounded theory [40]. This inductive approach to data was originally developed to counterbalance the deductive and positivist sociological method of validation and verification of existing theories, which in several cases proved insufficient when attempting to describe what was really going on in a certain sociological context [40].

The analysis of the transcribed dialogues happened through three levels of coding: open coding, selective coding, and theoretical coding [40]. In the process of open coding, detailed reflections from the dialogues as well as different topics that were touched upon through the dialogues were divided into a large number of subcategories. In the second coding process, the categories were merged into more superior categories, and at the same time, notes were taken on how the different categories appeared to be connected. All categories were subsequently examined in a third layer of coding through citations from the dialogues in order to understand the nature of the contexts better. This was also done to reach a conceptual understanding of the correlations that merged into theories and models describing the challenges and possibilities of combining hearing loss and work life [41].

The categories that represent reflections on the method, technology, and usability are to be found in Table 1. These categories have emerged through two levels of coding. Here, the categories concerning 'Role and identity', 'Control' and 'Biopsychosocial contexts' are greyed out as our focus in this paper is on the categories concerning the method in order to evaluate different aspects of this: 1. The process of data collection, 2. Data as an indicator for here-and-now experiences, and 3. Assistive perspectives of the method. In the following, the categories undergo a third layer of coding when looking into the participants' reflections related to the three categories.

Meta-categories			
Method /technology /usability	Role and identity	Control	Biopsychosocial contexts
1. The process of data-collection	*4. The role of the self in relation to others*	*6. To take control in life*	*11. Biopsychosocial contexts related to the hearing*
2. Data as an indicator for here-and-now experiences	*5. Self-perception*	*7. The relationship between feelings and control*	*12. Biopsychosocial contexts not directly related to the hearing*
3. Assistive perspectives of the method		*8. Maintaining control in the presence of other people*	*13. The importance of meaningful activities*
		9. Strategies as a means to maintain control	*14. What provides energy and what causes fatigue?*
		10. Dependency on assistive technologies	

Categories (label on left side of table)

Table 1-The three categories in the left column represent participant reflections on the method. The remaining categories are focused on matters that may cause or prevent the development of stress.

Results

The results will be presented as a summary of the main categories concerning the method identified in the coding of the data, accompanied by representative quotations from the participants.

The process of data collection

Some participants were excited about the technical aspects of the method. The gadget effect was particularly dominant among a group of the participants who were eager to get their hands on the devices, put on the HRV electrodes and find out how the ESM system worked. They did not show signs of nervousness and tried to fix technology breakdowns themselves during the week, either alone or with help from their spouses. One participant expressed a level of flow using the technology: "After the first day, I just started logging when the phone buzzed…I just logged how I felt at that particular moment." He seemed to stay at this level during the rest of the test week despite several technical problems, which for a short while sent him back to a level of mixed experiences and then allowed him to reach flow again after the problem was solved.

However, and in line with existing findings, we also experienced that technology breakdown can make people feel powerless [42]. Some participants were nervous about being alone with the system during the test week. They expressed insecurity and even anxiety about what to do if the system

behaved differently from what they expected or if a technology breakdown occurred. These observations are in line with studies on care technologies, where technologies that have been designed to provide care sometimes end up having the opposite effect making patients feel alienated and even causing anxiety and anger [42,43].

During the test week, the developers continuously updated the system. However, some updates decreased usability to such an extent that one participant became frustrated although she had started the week with great enthusiasm: "The latest update was quite bad" and "It is unsystematic when it crashes." In addition, the participant expressed her concern that the technical issues might have affected her mood and some of her answers during the experience logging.

A participant stated, "It took a little getting used to all the equipment, as people with hearing loss already carry around a whole lot of gear[1]. It was kind of a stress factor for me at least on the first day." This quick resolution indicated that some of the problems had to do with getting comfortable with using the system. However, in some cases, the technical challenges became too overwhelming and caused the participants to become frustrated or even unable to register experiences in the ESM. In several cases, they described the log-activity as a burden rather than a gift: "I couldn't turn it off during the night, and it was lying on the table buzzing until it fell on the floor… It was really annoying."

One of the two subjects who had to end the test prematurely because she felt that she was carrying too many technical devices around seemingly found it more difficult than another participant who carried around the same number of technical devices. The participant who ended the week prematurely had recently received cochlear implants and was getting used to the new hearing devices, and she was still in some pain after the operation. She was also in a stressful work situation as she was a trainee at a company as a part of her education, and the relationship with her immediate superior was not going well. This story indicates that there has to be some degree of stability in the lives of the participants for them to follow through with the test. In this case, the mindset depended highly on the resources available to the participant.

Data as an indicator for here-and-now experiences

Intensive data logging was necessary in the setup, as we wanted to be able to compare the detailed connection between the HRV and the momentary experiences. In addition, memory is potentially biased by time [44,45] and by logging experiences in the present moment, we believe that the participants remembered more precisely what had actually happened and what they felt at a particular moment. Logging the data did actually seem to help the participants remember - a participant stated, "The picture [the visual representation of

---

[1] Persons with hearing loss, who are fitted with hearing aids, cochlear implants or other hearing devices, often also have an FM (frequency modulation) system, which is a wireless sound transmission method to i.e. improve the sound from the television and from people talking.

the online data] fits very well with how I experienced the various situations."

All participants were curious to see how the HRV data reflected different activities. One participant was particularly curious to see whether feelings of frustration would show in his HRV data when he was engaged in a task that he felt was a waste of time. Another participant was curious to find out whether a specific situation, which drained her, would be traceable in the HRV data. There were also comments like "It's a lot of fun looking at your sleep pattern in your HRV". These comments indicate that the participants were enthusiastic when they looked at the HRV data, but the comments also tell us that the participants suspected certain situations to have an effect on not only their mental state but also their physical state. They were curious to have this suspicion confirmed or rejected through gaining an insight into their bodily reactions.

What seemed to be connected to the participants' motivation was whether they could relate to categories and questions in the ESM. There appeared to be a correlation between a clear understanding of the purpose and the relevance of the questions and an acceptance of the data showing 'the real me', which often led to constructive reflections on life circumstances and possible ways of creating change. In cases where the participants expressed an understanding of the categories and questions, they often interpreted the ESM-data as consistent with the physical and mental state they remembered to have experienced. One participant stated:

> In general, I can say that I have been confirmed in the feeling I had of my mood for the most part is good. And then I also think that the context I'm in, and the people I'm together with on a normal working day and in my spare time generally have a positive influence on my mental and physical balance.

Another participant stated:

> It's really nice to see how fast my body relaxes after such a conflict [...] I'm a bit surprised that I seem to be so relaxed while teaching because I often have to ask students to be calm and quiet. So, it's good to see that I'm still calm.

In both quotations, there seems to be a pattern: The participants interpret the measurements as something that can document a particular reaction or condition and thereby capture a here-and-now experience. The monitoring is not only used to "spot the self" but also to confirm "yes, that's right—this is actually how I feel". This experience of documentation seemed to be very motivating and led to further monitoring. The participants felt that the monitoring device was actually able to reflect their state of mental health and the situations and context they were a part of which created a meaningful setting.

In some cases, the participants had difficulties understanding the categories and questions. One stated, "I'm not sure how to answer the question of 'overview.' How should it be understood? Is it my own sense of overview, or is it my expectation of others' experience of my overview?" Another participant said, "I think the categories get abstract, and it's hard

to know how to answer the questions correctly." In both cases, the difficulties and reflections about how to answer the questions correctly had an impact on the participants' interpretation of the data from the monitoring. If doubt arose about whether the questions were answered correctly, the measurements were not considered a correct reflection of the condition of the participant. The participants experienced the measurements as demotivating, which led to a loss of belief in the devices and the ability to use the data.

Assistive perspectives of the method

Based on Antonovsky's salutogenic approach [30], each test course was initiated with an individual conversation, in which the test person set a personal goal or identified a particularly important topic, which could be a focus during the week of monitoring. All participants expressed a high degree of motivation for using the method to increase awareness of the processes and reactions in their body and mind, and it seemed very easy for the participants to identify something they were curious about regarding their own patterns of behavior and well-being.

A general reaction when the participants dealt with their own physical and mental measurements seemed to be that the participant positioned himself or herself outside his or her own experience and responded to it with a new experience, such as surprise. The monitoring became a way of seeing oneself through a second-order observer position [46]. A participant stated:

> I'm surprised that I had so much energy when I worked every night in the week leading up to the deadline Friday. On the other hand, I was completely exhausted Saturday. I have not really thought so much about how exhausted I am both physically and mentally after such a deadline is reached.

In this quote, the participant stated he was surprised by his own bodily and mental reactions to the workload he had experienced during the monitoring period. He acknowledged that the period had been stressful, but he became more aware of the degree of strain after he was confronted with the bodily and mental fluctuations the monitoring showed. He articulated that the self-monitoring data had been a new way to understand and reflect on how his work tasks and work hours affected him physically and mentally.

For this participant, the new insight became a positive motivational factor. He experienced empowerment, which led him to further desire to monitor. As he put it:

> Everyone should have access to this! I map many of the activities I do—use my calendar a lot. Here it is interesting for me to see how my body reacts, for example, when I work in the evening, go for a walk in the city or after an important deadline. Not everyone uses a calendar the way I do, and for those who don't, I think that this type of questionnaire is a great way to remember what you have been up to.

Other participants expressed the same feeling of empowerment and gain of new insight that made them stay motivated to monitor. Another example was a participant who feared a specific meeting with her manager, which, for her, was associated with major conflicts. Her mental reactions before the meeting were characterized by feeling depressed, having low energy and being generally unhappy about attending the meeting. Subsequently, the physical measurements showed that her body was surprisingly very calm during the meeting. She expressed her interpretation of the measurements as follows:

> I was excited to see how my body responded when I had to confront my manager with a problem I needed to talk to him about. I was surprised that I was so calm, and there was nothing at all to see at my heartbeat. It's a problem that has been haunting me for a long time, and the monitoring actually gives me energy to continue the struggle to get my manager to understand how I feel.

For this participant, seeing that her body apparently mastered the situation was an eye opener. This experience motivated her to continue the monitoring in order to reflect on how especially negative expectations could lower her energy level. She used the insight to work with her expectations in similar conflict situations, and she saw it as a reason to talk with her manager about specific situations in which she experienced a similar mental pressure.

However, not all participants had the same positive and motivating experiences when they looked at their physical and mental measurements. For some, the increased awareness seemed to be connected to concerns, which led to reluctance. A recurrent pattern among these participants was that they not only interpreted their current condition but also widely used the measurements to assess the risk of more negative potential health issues both physically and mentally. A participant articulated this topic as follows:

> I have a feeling that some things that the body absorbs and reacts to affect me. Because I can sometimes be so extremely tired—not physically but mentally… and I do not always understand why. I think that's because the body has reacted in situations that I wasn't really aware of and then I think: now that we've been wearing that heart rate monitor… is my heart also working overtime?

In this statement, the participant reflected on the question, "Is there something wrong with me that I have not been aware of?" The pattern of interpretation relates to a current condition but also to a potential negative condition. The increased awareness did not lead to sound thoughtfulness but to a kind of pathologizing where the risk of the potential undesirable condition invariably lurked around the corner. This way of understanding and interpreting the data seemed, for some of the participants, to lead to an understandable reluctance.

Conclusions

In this article, we have explored participants' experiences using the self-monitoring method Ecological Momentary Storytelling. The incentive to investigate the participants' experiences was, as mentioned initially in the article, an apparent gap between political ideology and the use of self-monitoring in practice. On the one hand, the use of self-monitoring is articulated as a way to empower the individual. On the other hand, it seems to be connected with difficulties to carry this empowerment out in practice. Health professionals experience how some patients resist and many projects using self-monitoring never become a long-term part of practice but stop at the pilot project stage. The majority of studies that deal with evaluation of self-monitoring efforts focus on either effects or economy. Our ambition with this study was to provide a qualitative contribution focusing on the participants' expressed experiences. We therefore asked the question:

Through which norms and interpretative frameworks do the participants understand the monitoring and its output? How is this interpretation related to either staying motivated or developing reluctance and/or dropping out?

Based on grounded theory, the participants verbalized experiences were divided into three main categories: 1) The process of data collection. 2) Data as an indicator for here-and-now experience. 3) Assistive perspectives of the method

There was apparently no correlation between the technical skills, which the participants possessed, and the ability to complete the test week – but technical challenges with the program crashing and the smartphone acting unexpectedly, clearly caused reluctance and even stronger feelings, like anger and irritation. This underlines the importance of interpreting the technology as reliable and as a relevant dialogue partner in order to stay motivated. The platform design should support and be flexible to a broad type of individual preferences and technical introduction and assistance should be available.

The findings also indicate that the interpretation of using the technology as a relevant dialogue partner is strongly connected to the experience of understanding categories and questions appearing on the mobile device. In cases where the participants expressed a clear understanding of the questions and categories, they also expressed confidence in the data showing a 'real me' and a true representation of their here-and-now experiences. This was clearly a motivational factor, which emphasize the importance of thorough introduction to basic concepts and underlying logics on the questions on the mobile device before the monitoring starts. Confusion about how to understand or answer specific questions often led to demotivation and lack of interest.

Regarding the participants' articulated experiences of interpreting the method as assistive, most of the participants regarded the method as a fruitful and helpful way of gaining and maintaining insight into own patterns of behavior and level of energy in connection with everyday life activities. They expressed how the new insight led to further motivation

for monitoring, because the data led to a new opportunity for changing existing negative patterns. In some cases, however, the insight did not lead to a positive awareness. Part of the participants interpreted the data with increased concern, wondering if the data showed any first sign of a negative development in their health and well-being. For some of the participants this increased concern led to declining motivation. For some of these participants motivation for monitoring was maintained, but in a way which appeared to be more pathologizing than empowering.

Discussion

Retrospectively, it has become clear to us how we, in the development of the method, were linked to an existing discourse on health and treatment. Establishing a methodological framework with inspiration from Antonovsky and Rogers, the focus was kept mainly at the individual and the individual's own opportunity to free up his potential and master the challenges he faced. In these methodological approaches, the basic assumption is that the individual is capable of coping, by increasing awareness of the management of energy in everyday life, and by gaining a meta perspective on both silent and spoken narratives. This means that the contexts and challenges anchored outside the individual is more or less absent. For example, we do not consider whether the test subjects actually face "unreasonable task overloading" or double-binding situations, which could be part of the focus in more context-oriented approaches to stressors [47].

It is important to emphasize that our study was not carried out as a part of the health care system and the participants were not diagnosed patients. Nevertheless, we believe that the above findings may help raise questions about the new forms of practice that emerge in the wake of the political headwind for self-governing treatment. As described initially, empowering the individual is a central political concept in the existing health promotion and treatment discourse. However, as the findings point out, being motivated or reluctant towards self-monitoring is not always anchored inside the individual. It is not only a matter of 'individual mastering', setting a goal or having a personal desire to be empowered. Based on the participants' expressed experiences being motivated or reluctant often reflects matters, which are anchored outside the individual; A complex life situation, experienced ability to handle and make technology fit into everyday life and the experienced resources to understand questions and categories posed on mobile devices. Another part of the findings shed light on how the increased awareness on body and mind seemed to be a double-edged sword leading both to experienced empowerment but also, for some of the participants, to increased concerns, which was pathologizing rather than empowering.

In continuation of these findings, it is relevant to address issues such as: How do we ensure that self-treatment takes into account complex life situations and varying individual resources? Moreover, how do we ensure that the individuals offered this form of treatment are those who can profit from it through experienced empowerment?

References

[1] Baker VH, Legen JE, Jacobsen CB. Patientinddragelse i Lægemiddlebehandling. København; 2016.

[2] Regeringen. Jo før, jo bedre - Tidlig diagnose, bedre behandling og flere gode leveår for alle [Internet]. København; 2014. Available from: http://www.sum.dk/~/media/Filer - Publikationer_i_pdf/2014/Sundhedsudspil-jo-foer-jo-bedre-aug-2014/Sundhedsudspil-Jo-foer-jo-bedre-aug-2014.ashx

[3] Kappelgaard LH. Vi skal passe på os selv - En diskursanalytisk undersøgelse af Sundhedsstyrelsens italesættelse af patienters selvmonitorering. Akad Kvarter. 2015;12:154–69.

[4] Frell M, Wandel A, Pedersen L, Jönsson ABR, Nyborg M. Læger og sygeplejerskers forståelse af patientinddragelse - En spørgeskemaundersøgelse blandt ansatte på hospitaler. København, Denmark; 2014.

[5] Pedersen KB, Nielsen SB, Land B. Behovet for teorier om sundhedsfremme. In: Dybbroe B, Land B, Nielsen SB, editors. Sundhedsfremme: et kritisk perspektiv. Frederiksberg: Frederiksberg: Samfundslitteratur; 2012. p. 27–39.

[6] Kappelgaard L. Det monitorerede mig - empowerment eller patologisering? Etikk I Praksis - Nord J Appl Ethics. 2015;9(1):37.54.

[7] Lupton D. The digitally engaged patient: Self-monitoring and self-care in the digital health era. Soc Theory Heal. 2013;

[8] Rasmussen JG. Inddragelse af patienter kan give store besparelser [Internet]. Dagens Medicin. [cited 2018 Jun 15]. Available from: https://dagensmedicin.dk/inddragelse-af-patienter-kan-give-store-besparelser/

[9] Sundhedsstyrelsen. Hvad mener behandlere om selvstyret behandling? Kvalitativ undersøgelse af holdninger blandt læger og sygeplejersker til kronisk syge patienters egenbehandling og selvmonitorering [Internet]. 2012. Available from: http://sundhedsstyrelsen.dk/da/sundhed/planlaegning-og-beredskab/specialeplanlaegning/~/media/2435A4D06FFF4BB787557F271D60F5C9.ashx

[10] Dalum AE, Arnfred S. Smartphone-apps som psykiatrisk behandlingstiltag. Ugeskr Læger. 2014;176(17):1577–80.

[11] Torpegaard H. Psykiater udvikler app mod panikangst. Dagnes Medicin [Internet]. Available from: https://dagensmedicin.dk/psykiater-udvikler-app-mod-panikangst

[12] Kappelgaard LH, Lund K. Ecological momentary

storytelling: Bringing down organizational stress through qualifying work life stories. Vol. 8027 LNAI, Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics). 2013.

[13] Kappelgaard L, Lund K. Er min hovedpine min hovedpine? -om selvmonitorering og refleksive dialoger som led i en sammenhængende og af-individualiserende forebyggelsesstrategi mod arbejdsrelateret stress. In: Frimann S, Sørensen MB, Wentzer HS, editors. Sammenhænge i sundhedskommunikation. 1st ed. Aalborh Universitetsforlag; 2015.

[14] Broens THF, Huis in't Veld RMHA, Vollenbroek-Hutten MMR, Hermens HJ, van Halteren AT, Nieuwenhuis LJM. Determinants of successful telemedicine implementations: A literature study. J Telemed Telecare. 2007;

[15] Potvin L, Balbo L. From a theory group to a theory book. In: Health and Modernity: The Role of Theory in Health Promotion. 2007. p. 6–11.

[16] Broendum E, Ulrik CS, Gregersen T, Hansen EF, Green A, Ringbaek T. Barriers for recruitment of patients with chronic obstructive pulmonary disease to a controlled telemedicine trial. Health Informatics J. 2016;24(2):216–24.

[17] Barello S, Triberti S, Graffigna G, Libreri C, Serino S, Hibbard J, et al. eHealth for patient engagement: A Systematic Review. Vol. 6, Frontiers in Psychology. 2016.

[18] Gambetti RC, Graffigna G. The concept of engagement A systematic analysis of the ongoing marketing debate Rossella. Int J Mark Res [Internet]. 2010;52(6):801–26. Available from: http://search.ebscohost.com/login.aspx?direct=true&db=bth&AN=56095859&site=ehost-live

[19] Frosch DL, Elwyn G. I believe, therefore I do. Journal of General Internal Medicine. 2011;26(1):2–4.

[20] Hibbard JH, Mahoney ER, Stock R, Tusler M. Do increases in patient activation result in improved self-management behaviors? Health Serv Res. 2007;42(4):1443–63.

[21] Middleton D, Brown SD. Social Psychology of Experience: Studies in Remembering and Forgetting. London GBR: SAGE Publications Inc. US. ProQuest ebrary. Web. 6 January 2015; 2005.

[22] Ptacek JT, Smith RE, Espe K, Raffety B. Limited correspondence between daily coping reports and restrospective coping recall. Psychol Assess. 1994;6(1):41–9.

[23] Raudaskoski P. Observationsmetoder (herunder videoobservation). In: Brinkmann S, Tanggaard L, editors. Kvalitative metoder en grundbog. 1st, 1 ed. 2010. p. 81–95.

[24] Shiffman S, Stone A a., Hufford MR. Ecological Momentary Assessment. Annu Rev Clin Psychol [Internet]. 2008 Apr [cited 2014 Jul 9];4(1):1–32. Available from: http://www.annualreviews.org/doi/abs/10.1146/annurev.clinpsy.3.022806.091415

[25] Stone a a, Schwartz JE, Neale JM, Shiffman S, Marco C a, Hickcox M, et al. A comparison of coping assessed by ecological momentary assessment and retrospective recall. J Pers Soc Psychol. 1998;74(6):1670–80.

[26] Wittgenstein L. Sidste skrifter om den filosofiske psykologi II: det indre og det ydre - 1949-1951. Århus: Klim; 2011.

[27] Hastrup K. Inkorporeret viden og praktisk kunnen. In: Perspektivet – kvalitativ forskning i arbejdsmiljø og arbejdsliv. Arbejdsmiljøinstituttet; 2005. p. 45--47.

[28] Sarah Pink and Kerstin Leder Mackley. Video and a Sense of the Invisible: Approaching Domestic Energy Consumption Through the Sensory Home. Sociol Res Online [Internet]. 2012;17(1):3 CHECK. Available from: http://www.socresonline.org.uk/17/1/3.html

[29] Zachariae B. Stress i et biopsykosocial perspektiv. In: Damsgaard-Srensen K, Madsen B, editors. Stress - Når kroppen siger fra. Gyldendals Akademiske Bogklub; 2003. p. 14–5.

[30] Antonovsky A. Health, stress and coping. Jossey-Bass; 1979. 255 p.

[31] Fredens K, Johnsen TJ, Thybo P. Sundhedsfremme i hverdagen. Fredens K, Johnsen TJ, Thybo P, editors. Munksgaard Danmark; 2011. 304 p.

[32] Kristiansen M, Bloch-Poulsen J. Kærlig rummelighed i dialoger - om interpersonel organisationskommunikation. Aalborg Universitetsforlag; 2000. 15--16 p.

[33] Rogers CR. Client-centered therapy. J Clin Psychol. 1951;7(3):294–5.

[34] Mccraty R, Ph D, Atkinson M, Tomasino D, Bradley RT. The Coherent Heart Heart – Brain Interactions , Psychophysiological Coherence , and the Emergence of System-Wide Order. 2009;5(2).

[35] Kristiansen J, Mathiesen L, Nielsen PK, Hansen AM, Shibuya H, Petersen HM, et al. Stress reactions to cognitively demanding tasks and open-plan office noise. Int Arch Occup Environ Health [Internet]. 2009 Apr [cited 2014 Oct 20];82(5):631–41. Available from: http://www.ncbi.nlm.nih.gov/pubmed/18936956

[36] Antonovsky A. Helbredets mysterium. Hans Reitzels Forlag, København; 2000.

[37] Annemans M, Audenhove C Van, Vermolen H, Heylighen A, Buildings R, Experiences R. Chapter 1 Hospital Reality from A Lying Perspective : Exploring a Sensory Research Approach. 2012;1:3–12.

[38] Strand AMC. On dis / continuous intra-active becoming of / through an Apparatus of Material Storytelling.

Aalborg University; 2012.

[39] Kappelgaard LH, Lund K. Foundations of Augmented Cognition [Internet]. Schmorrow DD, Fidopiastis CM, editors. Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics). Berlin, Heidelberg: Springer Berlin Heidelberg; 2013 [cited 2014 Nov 3]. 572-581 p. (Lecture Notes in Computer Science; vol. 8027). Available from: http://www.scopus.com/inward/record.url?eid=2-s2.0-84884822721&partnerID=tZOtx3y1

[40] Glaser BG, Strauss AL. The Discovery of Grounded Theory: Strategies for Qualitative Research [Internet]. Vol. 1, Observations. 1967. 271 p. Available from: http://www.amazon.com/dp/0202302601

[41] Lund K. Hearing loss, health, stress and work-life. Aalborg University; 2015.

[42] Hasse C, Wallace J. Omsorgsteknologier. In: Teknologi i sundhedspraksis [Internet]. 2014. p. 83–104. Available from: http://munksgaard.dk/Medicin/Teknologi-i-sundhedspraksis/9788762813427?gclid=CO_ix7Gr0sQC FazMtAodtlgAAg

[43] Hasse C. Artefacts that talk: Mediating technologies as multistable signs and tools. Subjectivity [Internet]. 2013;6(1):79–100. Available from: http://link.springer.com/10.1057/sub.2012.29

[44] Levine L, Safer M. Sources of bias in memore for emotions. Curr Dir Psychol Sci. 2002;11:169–73.

[45] Lalande KM, Bonanno GA. Retrospective memory bias for the frequency of potentially traumatic events: A prospective study. Psychol Trauma Theory, Res Pract Policy. 2011;3(2):165–70.

[46] Åkerstrøm Andersen N. Discursive analytical strategies. Underst Foucault, Koselleck, Laclau, Luhmann. 2003;

[47] Prætorius N. Stress - det moderne traume. 1st ed. Dansk Psykologisk Forlag; 2011.

Turning Points in Intermediate Patient Care Paths of Elderly: Constructive Reflections on Video Experiments with GPs and Municipalities

Helle Sofie Wentzer[a], Ann Bygholm[b]

[a] *VIVE - The Danish Centre for Social Research, Aarhus, Copenhagen, Denmark*
[b] *Department of Communication and Psychology, Aalborg University, Aalborg Denmark*

Abstract

The Danish healthcare system has transformed toward shorter hospital stays and increased dependency on primary care in municipalities. General practitioners (GPs) are key to preventing the (re)admission of elderly patients to the hospital, but visits to elderly, bedridden patients are not always compatible with GPs' office hours. This paper presents and discuss experiments with video in intermediate care paths for the elderly. The first experiment presents an ethical design guideline and playbook for cross-sectorial collaboration between a GP, home nurse and patient with video. The second experiment tries out video consultations with GPs in patient care paths. An ideographic, in-depth analysis of the communication and interaction between a 72-year-old male patient at a rehabilitation unit, his GP and municipality nurse give insights into clinical, organizational and technical aspects of video-mediated health care services. The analysis is reflected and discussed from a systemic perspective: At the micro-level, patient empowerment and safety in the patient care path from the video consultation are possible but cognitively demanding and risky for the role of the GP and the nurse. At the meso-level, interdisciplinary collaboration between the GP and the nurse depends on clarification of user roles, tasks and training in order for video to be efficient and safe. At the macro-level, the development of a cross-sectorial learning strategy, as well as a more thorough analysis of the kind of medical attention needed, is helpful for dividing tasks and responsibilities in intermediate care paths.

Keywords:

Video, Intermediate Care, Integrated Care, GP, Collaboration, User Involvement.

Introduction

Creating safe patient transitions from specialized treatment at hospitals to rehabilitation and care in municipalities is a well-known challenge for health care [1]. In the Danish context, two sectors deliver healthcare services: the primary sector of general practitioners (GPs) and municipalities (given the responsibilities of health prevention, care and rehabilitation) and the secondary sector of specialist care and treatment at hospitals. The patient care paths across sectorial boundaries, also called "the Bermuda triangle," are risky for patient safety and ineffective to marshal resources. Preventing the readmission of patients to the hospital shortly after they are discharged from the hospital, is a central concern to improve quality of care and patient safety, as well as to minimize the cost and resources for sending patients in and out of hospitals [2, 3, 4]. GPs play a pivotal role in the Danish healthcare system where they act as gatekeepers to secondary healthcare and manage most chronic and acute diseases [5]. GPs are central to prevent patients from being (re-)admitted to hospitals [6, 7]. The organization of general practice in patient consultations with 10- to 15-minute intervals between patients and a full waiting room leave little time for GPs to drive out on doctor's visits. Especially elderly, bedridden patients and patients older than 65 years with multi-morbidity and chronic diseases are most likely to be readmitted to the hospital [6]. Research shows that nursing-based case management and outgoing teams have no effect on readmission rates [8]. The GP is central in diagnosing ambulatory sensitive conditions. Intermediate patient care paths are challenged by the distance and workflow of general practice. There is a temporal and geographic distance to bridge in connecting the GP to the patient and the municipality nurse.

Materials and Methods

In a joint venture with the public, the region of Central Jutland, which runs the hospitals and pay the GPs, and the municipality of Aarhus arranged a 24-hour workshop with citizens to innovate patient care paths and the collaboration between the municipality and GPs [9].

Video was pointed out as a possible tool for supporting the communication and coordination of care paths across sectors especially when patients are discharged from the hospital to rehabilitation and care units in municipalities, and for virtual home visits for patients who require follow-up visits after hospitalization. The region has since collaborated with several municipalities, GPs and researchers to develop an infrastructure enabling video collaboration in the patient care

paths for elderly patients. This is part of the national policy on elderly patients and digitalization [10, 11, 12].

The aim of this research project is to collect context-sensitive knowledge on GPs' communication, interaction and collaboration with the patient and the municipality nurse via video consultations. The study's scientific background is ideographic (and not nomothetic); that is, the truth value relates to an interpretative phenomenological analysis of data, and not a quantitative analysis with a focus on frequency.

Data is generated from two experiments: The first experiment is a series of design workshops with an innovation group, representing different user and knowledge perspectives on integrated health care [13]. The second experiment is an ethnographic study of patient care paths to a rehabilitation unit (rehab unit) in the municipality for 24 hours of care after hospitalization [14]. Qualitative data is collected from observation, qualitative interviews [15, 16, 17], logging of data in the user-interface and video recordings of consultations [18]. An in-depth analysis of a patient case is presented to exemplify communicational aspects of patient safety in a video-mediated practice [19, 20, 21]. Permission to collect patient data and ethics was given by the Danish Authority [22].

The in-depth, phenomenological analysis is guided by the following questions: What takes place in the meeting, verbally and non-verbally (body language and examination)? How do the participants experience the clinical value of the intervention?

The results are reflected and discussed from a systemic perspective on the production of healthcare services. We differentiate between different levels of interaction:

- The micro-level, concerned with the direct interaction between the technology and the user, i.e., the video, the patient, the GP, the nurse (and others).
- The meso-level that refers to the professional and organizational context of use, i.e., the GP's office, the patient's home and the municipality's care setting.
- The macro-level considers the political and institutional system that frames the overall activities, i.e., rules, norms and work divisions.

Experiment 1: Design ethos and development of the playbook

An innovation group was established with participants from the two sectors and from the DaneAge Association, i.e., two GPs, two home nurses, two IT support specialists from the municipality, two IT supporters from the region and two elderly citizens.

The innovation group consented to a shared design ethos and developed a coordination and communication tool, "the

Playbook for Follow-up Home Visits by Video Consultation with GP" [13, appendix i].

Table 1 – Design ethos for cross-sectorial collaboration

Purpose of video-consultation	Efficiency, flexibility and safety/security for all participants
Participants	Patient, GP, Nurse and Providers (Health Care Law on Clinical Responsibilities of Patient Safety, Data Protection Law)

This principle guided the design process of the playbook, and served as the evaluation framework for testing the playbook in real-life settings. The playbook contains the procedures for interdisciplinary collaboration on video consultations between the GP and the municipality.

Playbook test scenarios

The playbook was tested in the private homes of the two elderly citizens participating in the project and at the clinics of the two GPs. Two patient scenarios were constructed as use cases, with the elderly participants acting as patients in their own homes. The two home nurses (who brought pillboxes as medication props) guided the video calls at each house and moderated the interaction between the patient and the GP. The two GPs were situated in their local clinics, behind their desks with a web cam installed on their PC screen. The IT support from the municipality and from the region helped the nurses get a Wi-Fi connection at the elderly individuals' homes and install the video clients on the nurses' laptops and on the GPs' PCs, including the speakers and webcam.

The tests showed that patient safety, efficiency and flexibility in collaboration depend on the following. See table 2.

Table 2 – Lessons learned from the testing of the playbook

Roles	Goals: Efficiency, Flexibility and Safety/Security
Video	A stable, Wi-Fi infrastructure and a software client with data encryption.
	The GP and nurse should be able to find each other easily as contacts in the interface. Phone numbers are exchanged as the second contact option.
Nurse	Book the patient's GP for a 30 min. video consultation.
	Prepare the clinical information on the patient's physical medication, prepare the elderly patient for communication via video with the GP, and arrange the camera/laptop in a convenient position to support the highest-quality sound and images.

	Call the GP at the scheduled time for the video consultation.
GP	Improved information in comparison with telephone calls from nurses trying to describe the patients' symptoms.
Elderly	Video consultation is desirable compared to no doctor visit.

The innovation group also pointed to the following risks: If the technical infrastructure fails, it has consequences for the whole health intervention, as the attention of the parties (the GP, nurse and patient) focused on establishing the Wi-Fi and software connection for communication. To avoid causing confusion, the video solution is not recommended for elderly patients with cognitive impairments.

Experiment 2: Video consultations in patient care path

Video consultations were tried out with patients in a rehabilitation unit, run by the municipality of South Djursland. The rehab unit has a stable Wi-Fi connection and 24 beds for patients, who have been discharged from hospitals but still require healthcare services. The unit offers 24-hour care and daily rehabilitation with an interdisciplinary team of a nurse, physiotherapist and occupational therapist. Twenty-four hours after a patient is discharged from the hospital, the patient's GP becomes responsible for diagnosis and treatment. Many of the patients' GPs live more than 18 km from the rehabilitation unit and are not obligated to conduct a doctor's visit. GPs from two clinics agreed to try out video as a possible medium of communicating with patients in the rehabilitation unit.

Video experiment with GPs

Five video consultations were carried out, with three GPs, four patients, a spouse and nurses from the rehab unit. For several reasons, the video consultations were very difficult to carry out. The reasons address organizational, technical and clinical perspectives.

Clinical perspectives

The GPs and the rehab unit agreed on sub-acute patients without dementia as possible participants in a video consultation. In practice, this was not a specific enough criterion for patient inclusion to secure a clear division of tasks and roles between the GP and the municipality nurse. Their communication was amplified by their separate and multiple contexts and therefore, cognitively demanding and emotional complex for the participants. On one side, the number of health issues that were addressed was impressive. On the other, communication and examination evolved erratically and unevenly, making it clear that all parties were uncertain about their roles and competences. The following analysis of a 27-minute recording of a video consultation between a GP, a patient and a municipality nurse gives some insights into some of the dynamics and challenges for video-mediated consultations between the GP and the municipality.

Themes of communication in video-mediated patient case

Five persons participated in the video consultation: the GP at his office and the 72-year-old patient and a nurse in the rehab unit together with the patient's younger sister and brother-in-law. A researcher observed and video-recorded the interaction in each setting. The analysis of the video recordings gives the following picture. The patient's medical story unfolds in the 27-minute video consultation with 11 themes:

i. The nurse, who initiates the communication between the patient and the GP, sums up the medical reason for the consultation, 'the patient's problem': The patient is paralyzed on one side after a stroke and doubts whether he should agree to be resuscitated in the event of a new stroke.

ii. The GP asks for the patient's blood pressure. The patient and the nurse confirm that it is stable.

iii. The GP takes a positive and motivational stance. He confirms to the patient, and indirectly to the relatives participating and the nurse, that he is still a strong man with a life ahead of him, and that he can benefit from training. He is "at the rehab unit because we believe in you."

iv. The GP observes tears on the patient's cheeks and asks if he gets any medication for depression. The nurse confirms he does and tells the doctor.

v. The sister sobs loudly during the consultation. She mentions her brother's future possibility of moving into a nursing home instead of going back to his private home.

vi. The nurse asks the doctor for correct treatment of a wound on his lower leg. The patient's sister shows a tube and asks whether the (ointment) medication can be used for the wound. It was prescribed by the GP's substitute before the stroke and hospitalization.

vii. The GP declines to give an opinion and recommends they arrange for a new doctor's visit during which a physical examination of the wound is possible.

viii. In the meantime, the nurse takes the camera and shows the GP close-up images of the leg ulcer.

ix. The GP changes his recommendation and says that use of the medication can continue.

x. The nurse explains how she will treat the wound, and the GP confirms her plan.

xi. The consultation ends by the GP repeating and emphasizing to the patient (and indirectly to the group) that he was discharged from the hospital to the rehab unit because they all believe in his will and strength to be able to regain some of his mobility.

User receptions of the video patient case

It appeared in the follow-up qualitative research interviews that the GP experienced a cognitive overload due to the many sources of verbal and non-verbal information over which he had little control. The nurse was uncertain about the technology, hardware and software, and depended on a

super-user to help her make the call and to activate the screen from the sudden stand-by mode. She was also uncertain about her role as moderator, relating to the patient and his sister and brother-in-law as their advocate and to the GP as a clinical supervisor to acknowledge her nursing competences. Thus, the GP and the nurse are put under pressure from the video consultation, with unclear benefits for their professional roles, and their ability to perform cross-sectorial and interdisciplinary collaboration.

The patient, however, experienced emotional relief from being visually perceived and acknowledged by his GP, especially because the GP knew him before the stroke, who he was as a healthy person, namely, very strong physically and mentally.

An existential situation

The video consultation had a positive impact on the patient's self-confidence and motivation to participate in the rehabilitation program. His interpretation of his own health situation changed from being "a vegetable not worth reviving" to consenting to be resuscitated in the case of a new stroke. In his GP, the patient had somebody who talked to him directly, and who sensed his "locked-up" situation: locked up from sitting in a wheelchair with one side of his body paralyzed but also the context of sitting between his brother-in-law and sister. The brother-in-law looked very uncomfortable with the whole situation, and his sister was crying, devastated by the patient's misfortune. His situation also put her in a new situation, practical as well as relational. She has to do his laundry (not part of the rehab unit's service). She lives several hours' drive away and was used to be in a sibling relation with him as her big brother, i.e., strong and protective. In contrast to this situation, the video presence of the GP offered the patient openings and possibilities of restoring his personhood and future life prospects.

Organizational perspectives

From an organizational perspective, different understandings of the importance of GPs to integrate care came to the forefront. The municipality has a general interest in strengthening collaborations with GPs in video consultations. Prevention of (re)hospitalization of the elderly saves resources in the municipality, and health prevention and rehabilitation are among the municipality's responsibilities. The municipality director of health prevention and home care has invested in the technical infrastructure of implementing Wi-Fi, computers, software, project management and education of super-users at the rehab unit. The region supports the development of an IT infrastructure and among others, the use of video in intermediate care. A special economic agreement is arranged with the GPs who participate in the experiment [23].

Ethical dilemmas in processes of intermediate care

The management at the rehab unit are ambitious to include medical resources as well. The geographic distance between the rehab unit and some of the GPs is a challenge, and driving patients to the GP or readmitting them to the hospital

happens frequently. The push from elderly patients coming in and out of the hospital to the rehab unit and back again, on occasions with patients dying a few hours after they have been resuscitated, put the care personnel in a dilemma. The procedure for asking patients to consent to resuscitation is part of this dilemma. Nurses are by the health care law obligated to start resuscitation efforts if patients suffer a heath attack, but for some of the frail patients, it might not always be the right thing to do; that is, dying, not life, is prolonged [24, 25]. In 2014, instructions from the Danish Health Care Authorities were given in order to make patients proactively decide on resuscitation matters. When patients decide against resuscitation, family members are to be included, and the GP, who has to confirm and document the patient's decision [26]. Asking patients about resuscitation then became part of the standard checklist at the rehab unit. The management has since renounced the procedure. As shown in the video analysis, the procedure also generates existential doubts in the patient's mind whether he or she deserves resuscitation. Thus, the procedure is partly responsible for creating the patient's ethical dilemma and "locked-up" situation. He is very dependent on the help of others, including his sister, and thus, is a burden, which makes it difficult for him to speak up for himself and insist on resuscitation in the event of a second stroke. Therefore, the GP becomes a great help to the patient. The GP, however, becomes part of a solution to a problem that is partly created by the accelerated care paths from the hospital to municipality care.

Resident doctor and consultant

The leader of the rehab unit favors more permanent access to medical help from a GP, preferably "two hours of daily visits" at the rehab unit. Resident doctors with frequent visits have been shown to increase the quality of the health care in nursing homes, among others, because the care staffs' competences increase [27]. Therefore, video is not the first choice, only a solution that is part of the daily medical issues.

Other care personnel pointed to the need for doctors outside GP office hours, as many patients' health conditions exacerbate at night, on the weekend and on holidays.

Competence in video-mediated collaboration

The leader of the rehab unit's nursing group is reluctant to push her nurses to perform new tasks with video, especially without sufficient training and IT support. The doctor-nurse collaboration is fragile in cases where medical tasks slide from doctors to nurses without adequate education. As more patients with more complicated needs are discharged, the pressure on municipalities' competences to secure patient safety has risen, and the dependence on collaboration with GPs has intensified. Interdisciplinary video consultation could push this already challenged collaboration too far in terms of the assignment of tasks, responsibilities and sufficient training.

Technical perspectives

The video plays an important role in creating a safe environment for the GP, patient and nurse to interact and communicate in.

Empowerment versus depowerment from video

From the patient's perspective, the video was an opening in his care path. He was given 'a voice' from the video consultation that was otherwise difficult for him to have on the subject of resuscitation. The communication with his GP is a positive turning point in the patient's care path. He is empowered to believe in the rehabilitation program, and in having a future. Nonetheless, the analysis points to an asymmetrical relation between the GP in his user context and the patient, nurse and relatives in their context of use, at the rehab unit.

Traditionally, in face-to-face communication, the asymmetrical relations between doctor and patient are interpreted as a power relation in favor of the doctor, reducing the role and influence of the patient to comply with doctor's orders. This is not the case in the video experiments for several reasons. The patient is empowered, but the GP, as well as the nurse, is depowered.

The GP has little control over the process of the patient interview. Questions comes from many, including the patient's sister, and make it difficult for the GP to predict and control the process on the difficult subject, which involves many emotions, but where the patient's needs have first priority.

The nurse is depowered because of the many responsibilities and tasks at the same time. She is responsible for the technical side but also for the patient, presenting his case, and toward the sister and brother-in-law. She is brought into a double-bind situation between the resuscitation issues and the many emotional responses and questions she has to moderate and care for.

Lack of video training

Another example of the depowerment of the GP and nurse involves a cancelled video consultation. The GP was waiting for the nurse's call, but the super-user at the rehab unit called in sick the same morning, and the nurse was not trained to do the call on her own. Thus, the GP waited in vain. Experiences of professional depowerment also affect the effectiveness of video, and make nurses as well as GPs more reluctant to schedule and agree to video consultations in the future.

Asymmetrical configuration of users in the video display

The depowering roles of the GP and nurse were partly caused by the constellation of the video camera. Because the camera angle has to cover four persons at a time, for the GP to see everybody, a distance is of approximately 3 meters is created between them and the camera and microphone.
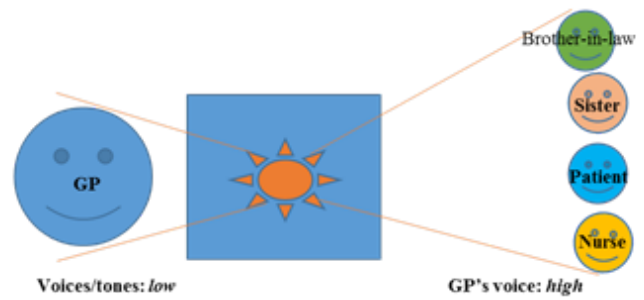
Figure 1- Asymmetrical configuration of video users

This distance has some disadvantages. The quality of the audio is not adequate. The GP cannot hear their voices properly, and repeatedly asks what they say (seven times. Visually, all four persons are displayed on the GP's interface in a diminished size. This means that the GP physically moves his chair and face as close to his PC screen as possible in order to get closer to see the facial expressions and body language better.

Another effect of this asymmetrical configuration is that because of the large size of the talking head everybody is figuratively addressed and encouraged to participate in the dialogue with the GP. Thus, instead of mainly the patient and the GP communicating with each other, the sister and the nurse also get very active in posing questions and setting the agenda of the interview. The diminished display of the group on the patient's side and the close-up face of the GP, therefore, demanded a lot of attention of the GP and of the nurse to compensate and respond in a "safe way" to the unpredictable process of the interview.

Results

Context-sensitive knowledge on video consultations in primary care is developed from the design ethos, the guideline for testing video consultations on elderly patients discharged from the hospital and the in-depth analysis of a video consultation in a patient care path. The analysis gave insights into clinical, organizational and technical aspects of the video consultation.

Context-sensitive knowledge

The clinical context is highly sensitive to ethical problems of patient care and treatment; thus, video consultations are to create safety for the patient, as well as to the GP and the nurse, who have professional responsibilities. The video also shares responsibility as a medium of communication and interaction. Encryption of video data is one aspect of safety/security. Another is the importance of symmetrical configuration of both contexts of use: the GP's office and the patient's and nurse's care setting. The configurations quality of sound and visual display are important features in order to align the relation between the communicators, i.e., the GP, patient and nurse. Family or other participants in video consultations, therefore, risk influencing the quality and efficiency of the interaction negatively.

From the organizational context, it appeared that collaboration with GPs is important to the municipality, but also that different forms of medical help are requested in intermediate care. Video consultation was only one out of three possible solutions that also related to different contexts and processes of care:

1. Video consultations with GPs to improve the individual patient care path
2. Consultancy from a GP who visits the rehab unit for 2 hours every day to do doctor's rounds and improve the quality of the treatment and care, including the competencies of the personnel
3. Emergency visits at nights and on weekends.

The collaboration between GPs and municipality nurses depends in general on adequate training and support in order for the doctor to delegate tasks and responsibilities to care personnel. In video-mediated collaboration, the user roles and tasks need to be defined and trained for the GP and the nurse to feel comfortable and professionally safe performing healthcare services that depend on video technology. The communication and interaction with video need more structure and well-defined tasks to be efficient and safe. Otherwise, the video consultations come at the risk of either being cognitively exhausting or do not take place as neither the nurse nor the GP is willing to take the trouble and risk.

Discussion

The results are not generalizable from a quantitative perspective on patient care paths, but qualitatively, the results contribute to the general understanding of the complexities in intermediate care, its ethical foundation and risks, but also of the possibilities for improvement. Within a systemic perspective, the results can be reflected at different levels of producing health care services. At the micro-level, patient empowerment and safety on the patient care path from video consultation is possible but cognitively demanding and risky for the role of the GP and the nurse. At the meso-level, interdisciplinary collaboration between the GP and the nurse depends on clarification of user roles, tasks and training in order for video to be efficient and safe. At the macro-level, the development of a cross-sectorial learning strategy, as well as a more thorough analysis of the kind of medical attention needed, is helpful for delegating tasks and responsibilities in intermediate care paths.

These reflections also emphasize the important recognition that video consultation does not solve the general need for GPs in municipality care units. There are more medical tasks, also related to doctors' rounds, supervision and consultancy, that do not fit into a universal solution of video. As for the use of video, it would be preferable to focus on continuing the improvement of the design of video consultations. The communication lacks structure in order to empower the GP. Delegation of tasks and planning ahead of the video consultation would give the nurse a better possibility of

performing her role, facilitating the communication and interaction between the GP and the patient with video.

To sum up, the design ethos of video consultations, i.e., to contribute to efficiency, flexibility and safety, depends on the symmetrical configuration of both user sides, and that the nurse and the GP are given the right competences to perform the specific tasks in the process: from scheduling the video consultation to preparing for it, and performing the patient interview and examination.

Conclusion

Video analysis can be a turning point in patient care paths for the elderly who require a doctor's visit. Patient empowerment is possible, but the interdisciplinary and cross-sectorial collaboration between GPs and municipality nurses is in a premature state. A cross-sectorial learning strategy needs to be developed based on a clearer definition of tasks and competencies related to video consultations. Not all medical tasks in intermediate care can be performed by video. Video consultations should focus on a structure that supports the individual patient's care path.

References

[1] DSR Analyse, NOTAT Flere korttidsindlæggelser blandt ældre patienter. January 2016 https://dsr.dk/sites/default/files/50/notat_flere_korttidsindlaeggelser_blandt_aeldre_patienter.pdf

[2] Defactum: Evaluering af fem indsatser i Region Midtjylland for den ældre medicinske patient. Region Midtjylland, Aarhus. Oktober 2016 http://www.defactum.dk/publikationer/ShowPublicationsAlphabetical?pageId=309986

[3] Vinge S, Buch MS, Kjellberg J, Kjellberg PK, Rasmussen SR, Wiuff MB, Forebyggelse af indlæggelser Oversigtsnotat om metoder, koncepter, evalueringer og effekter. DSI 2011 https://www.kora.dk/media/272122/dsi-3511.pdf

[4] Wiuff MB, Hansen EB, Kristensen TA, Madsen MH, Opfølgende hjemmebesøg i Region Hovedstaden. Kora 2017

[5] Hjertholm P, Clinical Activity in General Practice and Cancer. PhD dissertation, Faculty of Health Aarhus University 2015

[6] Purdy S, Avoiding hospital admissions. What does the research evidence say? King's Fund, 2010

[7] D'Souza S, Guptha S, Preventing admission of older people to hospital. BMJ 2013;346: f3186. doi: 10.1136/bmj. f3186.

[8] Audit Commission, More for less: are productivity and efficiency improving in the NHS? 2009

[9] Udviklingens døgn, Aarhus Kommune, Region Midtjylland praksis. Udvalgte projekter, bilag 2, 2011, Aarhus Kommune, Aarhus

[10] National handlingsplan, Styrket indsats for den ældre medicinske patient. Regeringen 2016, http://www.sum.dk/~/media/Filer%20-%20Publikationer_i_pdf/2016/Styrket-indsats-for-den-aeldre-medicinske-patient/National_Handlingsplan.pdf

[11] Nationale Strategi for telemedicine i Danmark 2012-2015. NSI Sundhedsstyrelsen 2011

[12] Strategien for digital sundhed 2018-2020. Sundheds- og Ældreministeriet, Finansministeriet og Danske Regioner, January 2018

[13] Wentzer, HS Opfølgende hjemmebesøg med video. Kora 2013

[14] Wentzer, HS Koordinering af teamsamarbejde i opfølgende hjemmebesøg med video – et eksperiment til innovation af tværsektorielle patientforløb. In: Sammenhænge i sundhedskommunikation. S Frimanm, M B. Sørensen & HS Wentzer, Aalborg Universitets Forlag, Aalborg 2015: 83-130

[15] Mol A, The Body Multiple: Ontology in Medical Practice. Duke University Press 2002

[16] Olesen F, Kroustrup J, AN T– Beskrivelsen af heterogene aktør-netværk. In: Introduktion til STS; Science, Technology, Society. Ed. CB Jensen et al. 2007: 63-91

[17] Kvale S, Interview. Introduktion til det kvalitative forskningsinterview. Hans Reitzels Forlag 1994

[18] Kushniruk A, Borycki E, Low-cost rapid usability engineering. Healthcare Quarterly 2006; 9(4): 98-100.

[19] Ricoeur P, Hermeneutics and the Human Sciences: Essays on Language, Action and Interpretation. Cambridge University Press 1981

[20] Ihde D, Expanding Hermeneutics – Visualism in Sciences. Northwestern University Press 1998

[21] Mattingly C, Moral selves and moral scenes: narrative experiments in everyday life. Ethos, International Journal of Anthropology 2012:301-327

[22] www.datatilsynet.dk

[23] Regionernes lønnings- og takstnævn, september 2017: Aftale om almen praksis. Danske Regioner, København, 2017

[24] Gadamer HG, The Enigma of Health. Stanford University Press 1996

[25] Kjeldsen SB, Ingen indikation for genoplivning. Sygeplejersken 2014;3:20-23. https://dsr.dk/sygeplejersken/arkiv/sy-nr-2014-3/ingen-indikation-for-genoplivning

[26] Sundhedsstyrelsen, Vejledning om fravalg af livsforlængende behandling, herunder genoplivningsforsøg, January 2014

[27] Weatherall CD, Lauritzen HH, Hansen AT, Termansen T, Evaluering af fast tilknyttede lægeordning på plejecenter. SFI, https://www.sst.dk/da/udgivelser/2014/~/media/DDDFF 580A76642FE8D95044073B39407.ashx (Access 01032018)

Address for correspondence

Helle Sofie Wentzer, Olof Palmes Alle 22,2. DK-8200 Aarhus N. Email: hewe@vive.dk

The Evolution of Clinicians' Preparedness for mHealth Use (2013-2017) and Current Barriers

Meghan Bradway[a,b], Lis Ribu[c], Gunnar Hartvigsen[d], Eirik Årsand[a,b]

[a]Norwegian Centre for E-health Research, University Hospital of North Norway (UNN),Tromsø, Norway
[b]Department of Clinical Medicine, UiT The Arctic University of Norway, Tromsø, Norwa,
[d]Department of Nursing and Health Promotion, Oslo Metropolitan University, Oslo, Norway
[c]Department of Computer Science, UiT The Arctic University of Norway, Tromsø, Norway

Abstract

Clinicians now insist that health authorities and researchers provide practical evidence and strategies for reacting to and handling patient-gathered data (PGD) and mobile health (mHealth) devices. With diabetes as a use-case, we present a summary of our own studies and a narrative scientific literature review to exemplify the progress of clinicians' perceptions of mHealth. We then compare these results to a narrative review of official clinical practice guidelines related to mHealth use (2013-2017) to demonstrate similarities and differences between what clinicians perceive as opportunities for mHealth and what health authorities are providing. Review of mHealth studies revealed that clinicians have become more willing to accept mHealth technologies and use patient-generated data over time. However, review of clinical practice guidelines revealed several barriers to using mHealth in clinical practice. Results of this comparison indicate 1) the need for a balance of clinician and patient participation and feedback during mHealth studies, and 2) health authorities' lack of sufficient guidance to clinicians for practically using mHealth in their daily practice.

Keywords:

Clinicians, Diabetes, Mobile Health, mHealth, mDiabetes, Consultation.

Introduction

Traditionally, medical devices for diabetes self-management and treatment were validated by health authorities. As such, clinicians were provided with structured guidelines and protocols for how to instruct their patients to use such technologies and relate to the subsequent gathered data. More and more commonly patient-operated mobile health (mHealth) tools enable patients to become more knowledgeable of their own health challenges and more in control of treatment priorities by providing them the means to better understand their own disease. As such, the novelties of mHealth throw a completely different spin on the priorities of patient care; clinicians are now expected to adapt not only to patients' new capacity to self-manage but also analyze larger patient-generated data sets.

Considering the lack of validation and testing within clinical settings, it is understandable that many medical personnel are concerned with various factors surrounding the clinical integration of e.g. mHealth apps [1]. Furthermore, because most often apps are designed for use by patients only, and not clinicians [2], initial evaluation studies within the medical realm focused upon answering questions relevant to individual patient users and not medical practice [3, 4]. Only until more recently was the concept of medical integration and evaluation considered [5]. Thus, medical personnel are now reacting to changes within two different environments: 1) the rapid increase of patient-centered mHealth, for example mobile diabetes (mDiabetes) tools, within the commercial sector as well as 2) pressures from patients to integrate such technologies within the medical sector.

The purpose of this paper is to identify the change in clinicians' perceptions related to mHealth between 2013 and 2017. By comparing this progress to the guidelines provided by regional and national health authorities, e.g. government agencies and those who create medical standards, we identify and emphasize the lack of necessary support for clinicians as well as the importance of including them in the planning and implementation of mHealth within clinical practices. This is especially important in primary health care, where research activities and partnerships with general practitioners' (GPs') offices are not as common as they are amongst health care personnel at hospitals.

Methods

Three narrative reviews were conducted. The first two were of health research literature, published between 2013 and 2017, that described mHealth interventions in which patient-gathered data were shared with clinicians. These were then compared to the third, which was a review of best practice recommendations produced by healthcare authorities, during the same period, regarding how clinicians should use patient-gathered mHealth data.

The first review was of mHealth interventions completed at our own University Hospital of North Norway's (UNN) Norwegian

Centre for E-health Research (NSE). These activities began with the REgioNs of Europe WorkING together for HEALTH (RENEWING HEALTH) Norwegian Pilot study (2013) [6, 7] in which individuals with Type 2 Diabetes were encouraged to discuss their use of an mHealth app for diabetes self-management, called the Few Touch Application (FTA), during consultations. During the 2014 annual Diabetes Research Conference in Oslo, Norway, we surveyed clinicians about their perceptions of a "clinician interface" of the patient-operated Diabetes Diary smartphone app. The next study that was conducted, concerning clinicians' use and relation to mHealth, was the Norwegian diabetes pilot of the international FI-STAR study [8]. Two GP's and a specialist participated in a clinician workshop in October 2016 to reflect upon what is needed to share patient-gathered mHealth data during consultations. In 2017 we invited patients and clinicians to participate in a co-design workshop, in both peer and joint sessions, to design their ideal mHealth data-sharing system, and indicate their preparedness for relating to mHealth. Workshops were audio-recorded, transcribed and translated into English.

To contextualize our own reports, we conducted a second narrative review of literature describing clinicians' perceptions of mHealth and patient-gathered data reported from similar studies within Europe and America. PubMed and Google Scholar were used to search scientific literature produced between 2013 and 2017. The following search strategy was used for PubMed: *clinician, practitioner, provider, or nurse AND barriers, concern, motivations, perspective, opinion, viewpoint or outlooks AND apps, mHealth, mobile health, wearables, or sensors*. The following search strategy was used for Google Scholar: *combinations of clinician, practitioner, provider, or nurse AND apps, mHealth, mobile health, wearables, or sensors*. Resulting articles were exported to the Systematic Reviews web app, Rayyan [9], for sorting and selection. Analysis included screening for reports of clinicians' firsthand experience with mHealth data presented by patients during clinical interventions. Author MB collected and reviewed the full-text of the publications, where data extraction included clinicians' perceptions of using mHealth in clinical practice, which were categorized as either perceived benefits or barriers. Benefits can be seen as clinicians' willingness to use mHealth, while concerns and needs represent the uncertainty toward using mHealth that needs to be resolved. Inclusion criteria were that literature must i) be published between 2013-2017 in English, ii) describe patient-operated mobile apps as part of the intervention, iii) describe studies included inquiry and reported responses of health care providers within America or Europe. Publications were excluded if i) they did not survey health care providers as part of the study, if ii) no abstract was found to support initial review processes, if iii) it only included "medical devices" [10], if iv) the intervention primarily provided basic mobile phone functions, e.g. SMS, from health providers for patient self-management.

- The third review was of official recommendations produced by health authorities related to how healthcare practitioners should react to, or use patients' own-gathered mHealth data or tools during consultations. Guidelines were searched for in European, Norwegian and American health authorities' websites including The World Health Organization (WHO), the European Commission (EU), Health Care Information and Management Systems Society (HIMSS), and the Norwegian Health Directorate, using versions of the following terms: *"clinical practice guidelines",*

recommendations AND Europe, Norway, America AND mHealth, mobile health, apps. Analysis included screening for any recommendations related to how clinicians themselves should react to and\or use patient-gathered data and mobile health technologies in daily clinical practice. This did not include recommendations for health facility managers or security systems. Data extraction included recommendations for how clinicians could relate to mHealth during consultations. These recommendations were then compared to the clinicians' needs to relate to, as presented in the previous two narrative reviews. Inclusion criteria were as follows: must mention daily medical activities performed by health professionals related to patient-operated mHealth technologies or their self-gathered data. Guidelines must also be published open-access between 2013-2017 within governmental, health authorities' and/or organizational reports. In focusing on publically available documents, we stress the importance of ease of access and use of these clinical practice guidelines for health care personnel themselves. Documents were disregarded if they i) provided no recommendations directly to health care practitioners for mHealth-use, ii) Only described design and/or evaluation guidelines for mHealth interventions studies, iii) Merely commented on issues related to mHealth-use during clinical practice, without direct input from clinicians themselves, iv) Only described appropriate use of clinicians' own mobile device during working hours.

Results

First, we summarized the clinician-related responses to mHealth interventions for our research activities, annually between 2013 and 2017. Table 1 summarizes the results related to clinicians' perceptions of the mHealth tools that were presented to them, both from previously published and unpublished (UP) reports from our studies.

Table 1- Own research: clinicians' perceptions related to use of mHealth tools and patient-gathered data (PGD)

Ref.	Benefits	Concerns and needs
2013		
UP	• PGD would be useful (n=17/23) • Would give recommendations based on PGD (n=17/23)	• Unclear financing (n=12/23) • Would require re-organizing services (n=11/23) • Training/supporting patients (n=11/23)
2014		
UP	• Better preparations of consultations (n=15/15) • Better able to help patients (n=13/15) • More effective communication with colleagues (n=9/15)	• More knowledge required about "patient compliance" (n=12/15) and "Integration into EHRs" (n=12/15) • Clinicians would need more "direct experience" (n=13/15) and training via "seminars" (n=13/15)
2015		
[11]	• Comfort with the system increase over time • Increase understanding of the patient situation • Graphical displays of data improved understanding of patient situation	• Not all patients present data, which is needed for clinicians to provide guidance

	2016	
[12]	• Easier to present PGD • Eager to discuss app data as graphs and trends • Patients reflect on data • Patients can and should take initiative during consults	• Must operate with existing medical technology • Data can be "noisy" • Patients need intensive training about how to collect data for medical purposes • Not all patients present data
	2017	
[13]	• Can base discussion and advice on personalized data • Result in more concrete discussions • Patients can become more engaged in their health • Specific information will save time	• Patients don't always present their data • Must be easy to collect data • Chance of data overload • Could be too time consuming • Clinicians still need to learn more about mHealth tools

Second, we summarize results of both narrative literature reviews of mHealth intervention studies and official clinical practice guidelines in order to contextualize our own findings and gain a greater understanding of the overall needs expressed by clinicians within the evolving field of mHealth.

The keyword searches in PubMed and Google Scholar results in 71 and 64 results, respectively. Initial review of titles and abstracts was based upon the inclusion criteria as described in the Methods section. 129 publications were excluded because medical practitioners were not directly surveyed and/or because the intervention did not involve patient-operated mHealth tools, leaving only 6 publications for full-text review (see Figure 1). Responses were then separated into benefits and barriers for relating to mHealth tools during clinical practice (see Table 2).
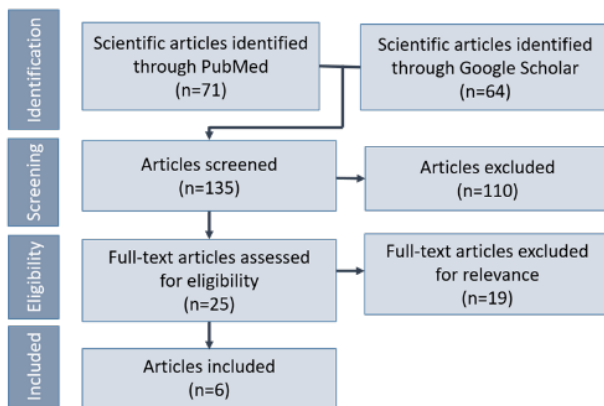
Figure 1- PRISMA flow diagram describing selection of scientific literature for review

Three of the reviewed publications reported clinicians' perspectives on the potential use of mHealth in general, while the remaining three papers reported clinicians' perspectives of a presented or tested mHealth system. The paper by Bonilla et al. [14] reported percentages of respondents' perceptions for each question, which allowed the authors of this paper to highlight how clinicians' perceptions differed between certain benefits and barriers. Table 2 summarizes the overall results of these six publications, ordered by publication year.

Table 2- Scientific literature search results: clinicians' perceptions related to use of mHealth tools and patient-gathered data (PGD).

Ref.	Benefits	Concerns and needs
	2013	
[15]	• Simple to use • Positive for patient care process • Monitoring patient progress • React to problems in real time	• Rapid technology progress • Unclear impacts of technology • Content/data overload • Too time consuming • May increase clinical workload • No operational support/guidelines
	2015	
[14]	• Tailored patient care • Supplementary patient support • Improved data accuracy • Increased amount of valuable data	• Cost • Time consuming to operate • Validity of applications • Reliability of patient-provided data • Risk of misinterpretation • Safety/privacy
[16]	• Tailor patient care • Improve communication • Fosters trust • Monitoring patient progress • Reinforce motivation and autonomy • Increase legitimacy of own clinical practice • Improve clinical workload/workflow	• Lack of direct view of patient data • Limited data flow and interoperability • Limited functionalities for follow-up • Technology limited to specific device • Competition with traditional care
	2016	
[17]	• More possibilities for teaching patients • Increase admin efficiency • Increase consultation efficiency • Supplementary patient support	• No operational support/guidelines • Limited data flow interop. • Time consuming • Tech. integration would compete with other priorities • No guidelines for handling sensitive information
[18]	• Would recommend apps • Comfortable exchanging info via technology • Monitoring patient progress • Improve communication • Must be endorsed by experts	• Discomfort using electronic communication with patient • Lack of sufficient evidence
	2017	
[19]	• Understand patient situations • Records symptoms • Medical adherence tracking & alerts	• No operational support/guidelines

Results from review of clinical practice recommendations

Ref.	Recommendations
2015	
[20]	Proposes Continua as the standard for welfare technology
[21]	**Guidelines for recommending apps to patients:** • Tailor app recommendations to patients and discuss consent regarding use of data and limits to consent • Discuss effective apps with colleagues • "Adhere to legislation and regulation (if existing) and/or professional obligations" • If the app is used for monitoring, the physician should instruct the patient how to respond to the information provided • Clinicians should look for the following characteristics before choosing an app: • Endorsement by professional or reputable health organization • Usability and evidence of impact - clinicians may also test the app themselves before recommending it • Reliability of information: inquire about how the patient intends to use the app to determine if the information provided is appropriate • Privacy/security: inform patients of added security risk of using apps, and even recommend apps with additional levels of authentication vs. apps without • Avoids conflicts of interest and fragmentation of health information
[22]	• Do not use medical apps that do not have a CE Mark, or if they do not "meet the requirements of the medical device directives and regulations" • "Exercise professional judgment before relying on information from an app"
[23]	• Clinicians should differentiate medical and non-medical mobile apps – differentiating characteristics are provided
2016	
[24]	• Clinicians should tailor recommendations to the disease and the mHealth apps/PGD presented by patients – example scenarios provided
2017	
	• None found

N=16 documents were identified from the search of clinical practice guidelines for mHealth. We excluded one document because it was not in English, and two documents because they were behind a pay-wall, leaving 13 for full-text review. We excluded 8 recommendations that do not offer practical solutions for clinicians in their every-day practice (see Figure 2).
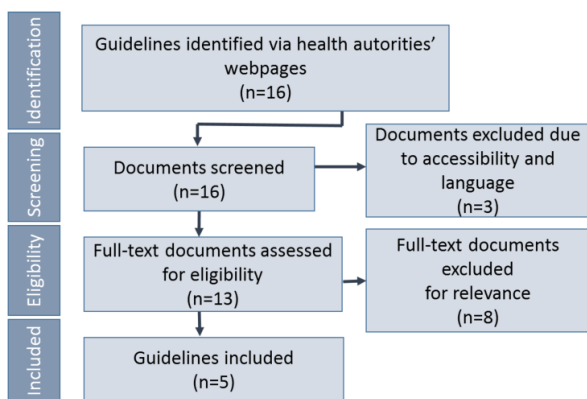

Figure 2 – PRISMA flow diagram describing selection of clinical practice guidelines for review.

Table 3 summarizes the recommendations provided by health authorities for how clinicians should relate to mHealth and patient-gathered data (PGD). This enabled us to compare if such recommendations meet clinicians' needs, as presented by the concerns and needs reported in Tables 1 and 2.

Table 3 – Health authorities' clinical practice recommendations for clinicians' use of mHealth and PGD.

Discussion

Clinicians have traditionally relied upon health authorities and management to provide guidance regarding clinical practice. As demonstrated, with the introduction of mobile health technologies to the options of patient self-management aids, clinicians have been and continue to be at a loss for answers. Despite these initial limitations, clinicians are acknowledging the benefits of these technologies more and more over recent years, especially since patients require more frequent support than the medical system is able to provide. Given the diversity of mHealth-generated data, health authorities and facility managers must provide support and suggestions for how care providers should relate to such technologies within differing clinical specialties in order for integration of mHealth to be successful.

The results of this paper also suggest an answer to the looming question; are the recommendations provided by regulatory bodies evolving quickly enough to meet the needs to clinicians in the rapidly changing environment of mHealth? Comparison of clinicians' perceptions of mHealth over time and guidance produced by regulatory bodies demonstrate that health and care authorities are beginning to propose the type of specific suggestions for relating to mHealth that clinicians need. However, the majority of the official activities under-way involve preparation for secure technological integration on the back-end. There have been few guidelines or recommendations for how clinicians can use data gathered by mHealth tools such as apps and sensors in daily practice. Questions remain regarding how patient-gathered lifestyle and health data should be weighted and considered along-side clinically generated information, e.g. lab results, to inform and generate actionable health recommendations. In addition, it is unclear which data is appropriate for providers to register and store within their own EHR systems. Health providers are responsible for judging which information is medically necessary and relevant for clinical decisions versus which information is sensitive to the individual and, therefore, should not be shared with the rest of the coordinated care team. This task is made exponentially more difficult with the added volume and detail of patient-gathered data, and our current research project *Full Flow of Health Data Between Patients and Health Care Systems* will address this in the coming clinical study of a mHealth system during clinical practice in Norway.

Conclusion

We have seen a development in mHealth where mobile technology, such as apps for mobile phones, smartwatches, and patient-operated sensors, have led to a situation in which patients are bringing new and more data into the clinical settings. mHealth is a rapidly developing field and clinicians need sufficient guidance to respond to the frequent changes and challenges that this new environment calls for. As this paper demonstrates, while official guidelines published by health authorities reference standards for back-end requirements for technological communication between EHRs and mHealth devices, they do not provide sufficient support for clinicians' in

their daily struggle to relate to mHealth. Therefore, the authors advocate for a greater voice and active involvement of health professionals in the development of any new processes, protocols or official standards, regardless of their specialty, to relate to mHealth successfully on a daily basis. It is time to integrate mHealth learning into medical and continued-education for practicing clinicians.

References

[1] Boulos MN, Brewer AC, Karimkhani C, Buller DB, Dellavalle RP. Mobile medical and health apps: state of the art, concerns, regulatory control and certification. Online journal of public health informatics [Internet]. 2014 June 15, 2015 [cited 2014; 5(3):[229 p.]. Available from: http://ojphi.org/ojs/index.php/ojphi/article/view/4814.

[2] Deign J. How Mobile Apps are Helping the Public Sector The Network [Internet]. 2015 June 30, 2015. Available from: http://newsroom.cisco.com/feature/1578416/How-Mobile-Apps-are-Helping-the-Public-Sector.

[3] Spook JE, Paulussen T, Kok G, Van Empelen P. Monitoring dietary intake and physical activity electronically: feasibility, usability, and ecological validity of a mobile-based Ecological Momentary Assessment tool. Journal of medical Internet research [Internet]. 2013 Pmc3785990]; 15(9):[e214 p.]. Available from: http://www.jmir.org/2013/9/e214/.

[4] El-Gayar O, Timsina P, Nawar N, Eid W. Mobile applications for diabetes self-management: status and potential. Journal of Diabetes Science & Technology. 2013;7(1):247-62.

[5] Eng DS, Lee JM. The promise and peril of mobile health applications for diabetes and endocrinology. Pediatric diabetes. 2013;14(4):231-8.

[6] Torbjornsen A, Jenum AK, Smastuen MC, Arsand E, Holmen H, Wahl AK, Ribu L. A Low-Intensity Mobile Health Intervention With and Without Health Counseling for Persons With Type 2 Diabetes, Part 1: Baseline and Short-Term Results From a Randomized Controlled Trial in the Norwegian Part of RENEWING HEALTH. JMIR MHealth and UHealth. 2014;2(4):e52.

[7] Holmen H, Torbjornsen A, Wahl AK, Jenum AK, Smastuen MC, Arsand E, Ribu L. A Mobile Health Intervention for Self-Management and Lifestyle Change for Persons With Type 2 Diabetes, Part 2: One-Year Results From the Norwegian Randomized Controlled Trial RENEWING HEALTH. JMIR Mhealth Uhealth. 2014;2(4):e57.

[8] Use Case Scenarios FI-STAR webpage: FI-STAR consortium; 2014 [Available from: https://www.fi-star.eu/use-cases.html.

[9] Ouzzani M, Hammady H, Fedorowicz Z, Elmagarmid A. Rayyan-a web and mobile app for systematic reviews. Systematic reviews. 2016;5(1):210.

[10] US Food and Drug Administration. Draft guidance for industry and Food and Drug Administration staff, mobile medical applications.

[11] Bradway M, Grøttland A, Blixgård HK, Giordanengo A, Årsand E. System for enabling clinicians to relate to a mobile health app: Preliminary results of the Norwegian trial in the EU FI-STAR project. Advanced Technologies & Treatments for Diabetes (ATTD). 2016.

[12] Bradway M, Holubova A, Joakimsen R, Årsand E. Differentiating Presentation Of Patient Gathered Data Between Type 1 And 2 Diabetes During Consultations. 10th International Conference on Advanced Technologies & Treatments for Diabetes; February 2017; Paris, France: Diabetes Technology & Therapeutics; 2017. p. A-1 - A-133.

[13] Bradway M, Giordanengo A, Grøttland A, Morris R, Årsand E. Design better together: Co-design workshop protocol to develop an m-diabetes data sharing system between patients and clinicians. Advanced Technologies & Treatments for Diabetes Conference 2018; 2018-02-14 - 2018-02-17; Wien2018.

[14] Bonilla C, Brauer P, Royall D, Keller H, Hanning RM, DiCenso A. Use of electronic dietary assessment tools in primary care: an interdisciplinary perspective. Journal of medical Internet research. 2015;15:14.

[15] Parmanto B, Pramana G, Yu DX, Fairman AD, Dicianno BE, McCue MP. iMHere: A Novel mHealth System for Supporting Self-Care in Management of Complex and Chronic Conditions. JMIR MHealth and UHealth. 2013;1(2):e10.

[16] Barnett J, Harricharan M, Fletcher D, Gilchrist B, Coughlan J. myPace: an integrative health platform for supporting weight loss and maintenance behaviors. IEEE Journal of Biomedical & Health Informatics. 2015;19(1):109-16.

[17] Mares M-L, Gustafson DH, Glass JE, Quanbeck A, McDowell H, McTavish F, Atwood AK, Marsch LA, Thomas C, Shah D. Implementing an mHealth system for substance use disorders in primary care: a mixed methods study of clinicians' initial expectations and first year experiences. BMC Medical Informatics and Decision Making. 2016;16(1):126.

[18] McClure JB, Hartzler AL, Catz SL. Design considerations for smoking cessation apps: feedback from nicotine dependence treatment providers and smokers. JMIR mHealth and uHealth. 2016;4(1).

[19] Simpson AJ, Honkoop PJ, Kennington E, Snoeck-Stroband JB, Smith I, East J, Coleman C, Caress A, Chung KF, Sont JK. Perspectives of patients and healthcare professionals on mHealth for asthma self-management. European Respiratory Journal. 2017;49(5):1601966.

[20] Solberg E, Høie B, Mæland M, Røe Isaksen T. The Government Action Plan for Implementation of the Health&Care21 Strategy Action Plan: Research and innovation in health and care (2015-2018). In: Ministeries N, editor.: Norwegian Ministry of Health and Care Services; 2015.

[21] Canadian Medical Association. Guiding Principles For Physicians Recommending Mobile Health Applications To Patients. Canada2015.

[22] Royal College of Physicians. Using apps in clinical practice: Important things that you need to know about apps and CE marking 2015.

[23] US Department of Health and Human Services. Mobile Medical Applications Guidance for Industry and Food and Drug Administration Staff. In: Administration FaD, editor. 2015.

[24] Norwegian Health Directorate. Attachment 2: Health care services track. 2016.

Address for correspondence

Meghan Bradway
Norwegian Centre for E-health Research
Postboks 35
9038 Tromsø
Email: meghan.bradway@ehealthresearch.no

Communication and Relations between Healthcare Professionals before and after Implementation of a Telehomecare System: A Study Protocol

Karsten Ulrik Niss

Department of Business and Management, Aalborg University, Aalborg, Denmark

Abstract

Twenty years after Reed Gartner stated, that a successful implementation of IT systems is 80% organization, and 20% technology, the organizational part of implementing health IT is still overseen or gives challenges in many projects.

This research project follows the implementation of a telehomecare project for citizens with heart failure and focuses on the communication and relations between the health professionals.

The hypothesis is that new forms of cooperation and relationships between the actors emerge, mediated by the telehomecare technology used. The methodology includes Actor-Network Theory (ANT) and Relational Coordination as the conceptual framework. Data is collected using a mix-methods approach combining a before/after survey and focus group interviews. The relations between the actors are analysed using Social Network Analysis (SNA), and data from the interviews are analysed using an iterative condensation.

The objective is to show that relational coordination, ANT, and SNA, are useful theories and tools to elucidate changes in collaboration before/after implementation of a telehomecare solution, and provide an indication of why implementation in comparable organizations gives different results.

Keywords:

Telemedicine, Medical Informatics, Organisations, Organisational Change, Social Network Analysis, Actor-Network Theory, Relational Coordination.

Introduction

It is now twenty years ago that Reed Gardner, at a health informatic conference in 1998, stated that successful implementation of IT systems is 80% dependent on the development of the social and political cooperation skills, and 20 percent or less on the hardware and software implementation [1, 2]. Marc Berg [3] supported this view by claiming that technology is crucial,

but secondary – even today several IT projects in the healthcare sector do not prioritize the organizational element of implementing IT enough.

In addition to the fact that the Danish municipal reform in January 2007 reduced the number of municipalities from 271 to 98, and the 14 counties became five regions, municipalities became more important in the health sector [4]. The municipalities gained major responsibility for prevention efforts, care and rehabilitation in relation to citizens, and the municipalities should contribute to the financing of health care in the future [4]. The latter was in order to encourage the municipalities to effectively carry out the tasks in the field of health. The changed division of tasks between primary and secondary sectors calls for closer cooperation with an increased need for coordination of treatment and care of patients/citizens between the sectors. To support the closer cooperation, new IT solutions have been developed for the exchange of information between the sectors – among these, an increased focus on telehomecare.

The use of telehomecare in general is growing and many local anchored pilot projects have been implemented, of which few have become decisive operating projects - with even fewer becoming large-scale or nationwide projects. Between 2012 and 2015 TeleCare Nord[1] ran a large-scale project in the Northern Jutland Region, Denmark. The purpose of the project was to offer telehomecare to citizens with Chronic Obstructive Pulmonary Disease (COPD). The experience of this project (hereinafter referred to as the COPD project) is part of the foundation for the rollout of telehomecare for citizens with COPD at national level. At the same time, the experience from the COPD project is being used for a new large-scale project, where telehomecare is to be offered to citizens with heart failure - in daily terms called the Heart Failure Project or just Heart Failure.

In the TeleCare Nord COPD project, there has been a strong focus on the organizational issues. Some of the experiences from the COPD project are, that "…it is an intensive work to adjust and adapt services and procedures to get coherent and effective processes." [5]. In spite of the above focus on the organizational issues, it is stated that "... should remain operational for both solutions and cooperation relationships." [5] and

---

[1] Established in 2015 as a cooperation between the Northern Jutland Region and 11 municipalities in the region to support telehomecare projects.

in the completion report for the COPD project, the importance of "... a constructive and close cooperation across the core actors." [5], and that cooperation is not "... implemented to the bottom, and there must be continued focus on behavioural- and practice changes." [5]. Another experience from the COPD project is that even with the same technology, the implementation has been more successful (measured by activity between the healthcare professionals – the actors) in some places, than in other places.

Purpose

The purpose of the research project is to do a descriptive/exploratory study [6, 7] to:

- Elucidate cooperation relationships between the actors before and after the implementation of TeleCare Nord Heart Failure to describe the importance of implementing TeleCare Nord Heart Failure on collaborative relationships.

- Elucidate whether there is a correlation between the cooperation relations and the outcome of the implementation, depending on where the actors come from.

- Understand and explain why the opinions and identities of the actors in terms of cooperation/relative coordination are changed/not changed before/after the implementation of TeleCare Nord Heart Failure.

- Understand and explain why there is a/is no correlation between the cooperation relations and the outcome of the implementation, depending on where the participants come from.

The working assumptions for the research work are:

- The implementation of TeleCare Nord Heart Failure leads to a higher degree of collaborative relationships/relational coordination, as the functionality of the application supports a greater degree of cooperation between the different actors.

- That the cooperation relations have a major impact on the outcome of the implementation of a telehomecare system - and that close collaborative relationships in an organization lead to better implementation than organizations with more loyal and sporadic cooperation relationships.

The purpose is chosen based on the evaluation of the COPD project and has been done in cooperation with TeleCare Nord. It is the intention that this PhD Project can help generate knowledge, that can subsequently help minimize the organizational challenges that emerged from the COPD project, as mentioned earlier, and thus be of benefit to other similar telehomecare projects.

The scope of the project is to contribute knowledge about communication/relations between actors through the implementation of telehomecare solutions. The currently included actors are:

- The health professionals at the four hospitals in the Northern Jutland Region

- Care staff in the 11 municipalities in Northern Jutland

- General practitioners in the Region of Northern Jutland

However, the citizens with heart failure, as an actor, have been excluded from the study. This is done to keep the focus on the organizational conditions and not the individual's medical progress.

The scientific purpose is to be able to use the theory of relational coordination and Actor Network Theory (ANT) as the conceptual framework for:

- Understanding change in relationships between health care professionals before/after the implementation of the telehomecare solution.

- Understanding why there is a difference in the implementation depth of the different groups of health care professionals.

- Suggesting one solution to what the healthcare sector can do to optimize the implementation of telehomecare solutions.

Materials and Methods

The overall methodological approach is to use relational coordination as the theoretical framework, ANT as the conceptual framework and Social Network Analysis (SNA) as the data collection and analysis tool, but also to incorporate other relevant theories to discuss the outcome of the above.

Conceptual framework

The conceptual framework is based on the assumption that relational coordination is based on two "legs" – (1) relationship between the actors, and (2) coordination task between the actors. The focus of this project will be on the relationships, and to describe relations between the actors as well as between the actors and the technology – telehomecare. ANT is included, as ANT focuses on networks consisting of heterogeneous actors, things, facts, etc. Gittell's relational coordination will be discussed in regards to other management literature, which is not currently selected – but could be related to change management [8], Balogun & Hailey's cultural web [9], and/or wayfinding/wayfaring [10-13].

As the general theoretical framework for analysing and discussing cooperation relationships between the participants in the Heart Failure project, Jody Gittell's [14-16] concept of "relational coordination" is used. Relational coordination has been used as inspiration in building the former and current organizational structure of the Northern Jutland health service as well as across the sectors. Jody Gittell's concept of relational coordination is used to understand what skills (professional, personal and social), that are required for collaboration in the Heart Failure project to work. ANT is involved as the conceptual framework for interpreting and analysing the relationship between the individual health professionals, and between them and the technology used (telehomecare). This involves both an understanding of why actors do what they do, and the perception, that the actors attach to why they do as they do. In this way, an

understanding of the interaction between professional identity, technologies, and other material circumstances, that enables or conversely prevents successful implementation of the Heart Failure project, is created.

As a tool for analysing data about actors' collaborative relationships, SNA is used. SNA is a structured approach to uncovering and describing networks [17-20]. By using SNA, it is expected that an insight can be given to the difference between communication and relations between the actors in the individual municipalities and to show whether there has been a change following the implementation of Heart Failure. The data collection for the SNA will be via electronic questionnaires and interviews with health professionals.

Paradigmatic background

The project forms part of a theoretical framework regarding the development of management technologies and organizational structures used in public administration. Overall, this has been a development from New Public Management (NPM) towards New Public Governance (NPG). New Public Management was introduced by Christopher Hood in an article from 1991 [21], but since the 1980s it has been the governance paradigm in the public sector, where efficiency, performance management and continuous change are the goal and governance through incentive management [22, 23].

In addition, NPM works from a market mechanism and customer orientation where the patient/citizen becomes users [22]. As a replacement – or perhaps more in addition to the NPM, there has been more focus in recent years on NPG, where collaboration in network/partnership and innovation is in focus. This is because, among other things, NPG's focus: "... is very much about inter-organizational relationships and the governance of processes, and it stresses service effectiveness and outcomes." [24]. The NPG is based on a fundamental assumption, where the public sector is seen as an arena of co-operation and where network management and innovation are in focus [25].

At the moment, I see that relational coordination can be used both within a governance paradigm based on NPM, but also - and perhaps most – within a governance paradigm based on NPG.

Relational Coordination, ANT and SNA

Relational coordination is largely based on network thinking [15], where Gittell uses some of the same calculations (eg Strengths and Cronbach's alpha) and chart types, which are also used in SNA. Until now, however, no explicitly reference to SNA has been in the read litterature, but it is my view, that the use of SNA, as a tool for analysing relational networks, will be obvious.

One of the possibilities of combining ANT and SNA is that it will be possible to work with a larger number of actors, and that it will be possible to follow them over time as data collection to SNA can be done via quantitative methods where ANT primarily works with qualitative approaches [26]. This is supported by Latour et al. [27], which discusses the combination of SNA and ANT, and among other things, highlights SNA's ability to display and follow actor networks over time.

Venturini et al. [28] concludes that there are important common features of SNA, ANT, and digital networks that allow (SNA's) graphs to be used in the study of actor networks, although there are many differences.

At the same time, I see that ANT can be used to discuss the thinking behind relational coordination and thus a logical relationship between relational coordination, ANT and SNA.

Data

The project includes a before/after study, with the purpose of seeing if the implementation of Heart Failure causes a change in the cooperation, relationships and communication between the involved actors. Later a study based on focus group interviews is planned. The purpose is to create a dataset, that makes it possible to compare the use and implementation of Heart Failure across the 11 municipalities. In addition to providing material from the municipalities, the interviews also provide material for comparing the data collected in the before/after study and the study based on interview as a data collection method.

Before/after study

Data collection is based on an "adapted prospective panel longitudinal" study [29]. "Adapted" as data collection first occurs after some of the respondents have begun to use or have been trained to use the Heart Failure solution. This could speak for calling the first data collection retrospectively, but since other respondents at the data collection point had not begun to use or trained in the Heart Failure solution, it is decided to call it the "adapted" approach. The "prospective panel longitudinal" is chosen, as the study follows the same population over time before implementation, and has two data collection points over the period [29, 30].

The pre-survey was conducted as a questionnaire survey and was done via an electronic questionnaire in SurveyXact. Access to the questionnaire was via link in a mail sent to the respondents. The questionnaire has been sent to the 11 municipalities, four hospital locations and 175 GPs in the Region of Northern Jutland. Three reminders from the TeleCare Nord secretariat were sent to the respondents. In order to be stringent, the questionnaire is used again for data collection for post-analysis, scheduled to take place in September 2018, when Heart Failure has been in operation for some months.

After examination

In addition to the questions relevant to the SNA analysis, interviews will include questions that implicitly address relational coordination, ie. relating to common goals, shared knowledge, and mutual respect, as well as questions that implicitly address the technology-actor in relation to ANT.

After the last data collection and analysis, it is planned to have a focus group interview to go into depth with the result and discuss possible explanations of these with selected respondents. The combination of quantitative and qualitative study will be done with inspiration from mix-methods [31, 32] and successive triangulation [33], to seek an explanation of the results of the quantitative examination.

The goal is to understand and possibly explain why there is/is no correlation between the relationship of cooperation and the outcome of the implementation, depending on where the participants come from.

SNA as an analysis tool

The purpose of using SNA in this project is to use a recognized approach to analyse the communication/relationships between the health professionals in the Heart Failure project, thereby creating the basis for working with ANT - and ultimately relational coordination. The aim is to elucidate and, if so, to what extent the number of relationships between health professionals' changes after the implementation of Heart Failure. The limitation of looking at the health professionals is justified by the fact that the organizational experience from the COPD project does not include the citizens with COPD, but points to organizational conditions among and relationships between the health professionals.

Implications

The understanding before starting the study is that patients were previously in physical contact with the hospitals and the GP's, but that the care staff via the telehomecare solution have the opportunity to quickly and easily answer questions - either from the GP or from health professionals at one of the hospitals. Previously, the citizen had to meet physically at the hospital or at the GP to make the measurements that it is currently possible to make in the citizen's home via the Heart Failure solution. At the same time, the GP is able to confer measurements from the patient with the health professionals at one of the hospitals in an easy and fast way through the Heart Failure solution.

After the implementation of the Heart Failure solution, the hypothesis is, that new forms of cooperation and relations between the actors are seen, mediated by the telehomecare technology used.

The objective is to show that relational coordination, ANT, and SNA are useful theories and tools to elucidate changes in collaboration before/after implementation of a telehomecare solution and can be used to provide an indication of why implementation of the same technology in comparable organizations gives different results.

On the empirical side, the goal is to give TeleCare Nord an insight into how relations and collaboration change after implementing a telehomecare solution, as well as through the final discussion of the results, to chart why the collaboration is changed/not changed - and perhaps suggest, what can be done to change the cooperation. This is in anticipation of providing knowledge that can be used to address and possibly minimize the organizational challenges in relation to the relationships and collaboration between the health professional and across the sectors. In addition, if a difference can be seen between the different organizational units.

Acknowledgments

References

[1] Lorenzi NM, Riley RT. Managing Change: An Overwiev. Journal of the American Medical Informatics Association. 2000;2000(7):116-24.

[2] Ball MJ, Douglas JV. Human Factors: Changing Systems, Changing Behaviors. In: Demetriades JE, Christophersen GA, Kolodner RM, editors. Person-Centered Health Records. Health Informatics Series. New York: Springer; 2005. p. 60-70.

[3] Berg M. Introduction. In: Berg M, editor. Health information management - Integrating Information Technology in Health Care Work. Routledge Health Management. London: Routledge; 2004. p. 1-8.

[4] Sundhedsministeriet I-o. Kommunalreformen - Kort fortalt. In: Indenrigs- og Sundhedsministeriet øk, editor. København: Indenrigs- og Sundhedsministeriet; 2005. p. 55.

[5] TeleCareNord. TeleCare Nord Afslutningsrapport - telemedicinsk storskalaforsøg i Nordjylland. Aalborg: Region Nordjylland, 2015.

[6] Andersen I. Den skinbarlige virkelighed - om vidensproduktion inden for samfundsvidenskaberne. 5 ed. Frederiksberg C.: Samfundslitteratur; 2013.

[7] de Vaus DA. Surveys in Social Research. 5 ed. Bulmer M, editor. London: Routledge; 2002.

[8] Jacobsen DI. Organisasjonsendringer og endringsledelse. 2 ed: Fagbokforlaget; 2012.

[9] Balugun J, Hailey VH. Exploring Strategic Change. 3 ed: Prentice Hall; 2008.

[10] Chia R, Holt R. Strategy as Wayfinding. Paper for the 24th EGOS Collogium, Amsterdam, June 2008

[11] Larsen ML, Rasmussen JG. Strategisk ledelse som meningsskabende processer. Jurist- og Økonomiforbundets Forlag; 2013

[12] Mintzberg H, Waters JA. Of Strategies, Deliberate and Emergent", Strategic Management Journal, 1985;6:257 – 72.

[13] Rouleau L, Balogun J. Middle Managers Strategic Sensemaking and Discursive Competences. Journal of Management Studies 2011;48(5):953-983.

[14] Gittell JH. Effektivitet i sundhedsvæsenet - samarbejde, fleksibilitet og kvalitet. 1 ed. København: Munksgaard; 2012.

[15] Gittell JH. Organizing work to support relational coordination. International Journal of Human Resource Management. 2000;11(3):517-39.

[16] Gittell JH. Relationers betydning for høj effektivitet. Styrken ved relationel koordinering. København: Dansk Psykologisk Forlag; 2016.

[17] Waldstrøm C, Engelbrecht B. Ledelse af netværk - Virksomhedens skjulte ressource. 2 ed: L&R Business; 2013.

[18] Provan KG, Milward HB. Do Networks Really Work? A Framework for Evaluating Public-Sector Organizational

Networks. Public Administration Review. 2001;61(4):414-23.

[19] Pryke S. Social network analysis as a research method. In: Pryke S, editor. Social Network Analysis in Construction. 1. 1 ed. London: Blackwell Publishing Ltd.; 2012. p. 77-100.

[20] Carrington PJ, Scott J. Introduction. 2014 2016/12/14. In: The SAGE Handbook of Social Network Analysis [Internet]. London: SAGE Publications Ltd; [1-8]. Available from: http://sk.sagepub.com/reference/the-sage-handbook-of-social-network-analysis.

[21] Hood C. A Public Management for All Seasons? Public Administration. 1991;69(1):3-19.

[22] Lerborg L. Styringsparadigmer i den offentlige sektor. 2 ed. København: Jurist- og Økonomforbundets Forlag; 2011.

[23] Greve C. NPM eller/og NPG - Hvad kan de bruges til? København: Væksthus for Ledelse; 2012 [cited 2017 6/12]. Available from: https://www.lederweb.dk/strategi/organisationsudvikling/artikel/100867/carste-greve.

[24] Osborne SP. The New Public Governance? Public Management Review. 2006 6/12-2017;8(3):377-87.

[25] Torfing J, Triantafillou P. Introduktion: New Public Governance på dansk. In: Torfing J, Triantafillou P, editors. New Public Governance på dansk. København: Akademisk forlag; 2017. p. 7-40.

[26] MacNeil RT, Mills AJ. Organizing a precarious black box: An actor-network account of the Atlantic Schools of Business, 1980–2006. Canadian Journal of Administrative Sciences / Revue Canadienne des Sciences de l'Administration. 2015;32(3):203-13.

[27] Latour B, Jensen P, Venturini T, Grauwin S, Boullier D. 'The whole is always smaller than its parts' – a digital test of Gabriel Tardes' monads. The British Journal of Sociology. 2012;63(4):590-615.

[28] Venturini T, Munk A, Jacomy M. Actor-Network VS Network Analysis VS Digital Networks. Are We Talking About the Same Networks? http://www.tommasoventurini.it/wp/wp-content/uploads/2015/05/Venturini-Munk_Jacomy_ANT-vs-SNA-vs-NET.pdf2015 [cited 2017 22/11]. Available from: http://www.tommasoventurini.it/wp/wp-content/uploads/2015/05/Venturini-Munk_Jacomy_ANT-vs-SNA-vs-NET.pdf.

[29] de Vaus DA. Research Design in Social Research. 1 ed. London: Sage Publications Ltd.; 2001 2001.

[30] Bowling A. Research Methods in Health - Investigating Health and Health Services. Buckingham: Open University Press; 1997 1997.

[31] Curry LA, Krumholz HM, O'Cathain A, Clark VLP, Cherlin E, Bradley EH. Mixed Methods in Biomedical and Health Services Research. Circulation: Cardiovascular Quality and Outcomes. 2013 10/1-2017;6(1):119-23.

[32] O'Cathain A, Murphy E, J. N. The quality of mixed methods studies in health services research. J Health Serv Res Policy. 2008 12/1-2017;13(2):92-8.

[33] Hollensen S. Triangulering. Et redskab i analyse- og dataindsamlingsprocessen. In: Darmer P, Freytag PV, editors. Erhvervsøkonomisk undersøgelsesmetodik. 2 ed. København: Samfundslitteratur; 1996. p. 227-38.

Address for correspondence

Karsten Niss, Department of Business and Management, Aalborg University, Aalborg, Denmark, E-Mail: niss@business.aau.dk

Usability of Eye Tracking for Studying the Benefits of E-learning Tutorials on Safe Moving and Handling Techniques.

Mette Hornbæk, Julie Hellevik, Clara Schaarup, Mette Ole Hejlesen

Department of Health Science and Technology, Aalborg University, Denmark

Abstract

Eye tracking is a measurement technology that quantifies the movement of eyes on a motive of interest, usually a screen. The technology has not before been used to evaluate the effect on e-learning used for teaching staff safe moving and handling techniques for moving patients. Based on four participants, we explored whether eye tracking could be used in combination with observations and interviews to assess the practical skills obtained after watching videos from an e-learning tutorial, teaching safe moving and handling techniques. The participants reported to obtain the most knowledge from what they saw in the video, rather than from what they heard or read. However, there was no clear correlation between time spent looking at Areas of Interests (AOI) in the videos and how the participants performed the safe moving and handling technique afterwards. Still eye tracking has potential as a measurement technology for providing objective knowledge that can be used to support qualitative data on the performance of practical skills in safe moving and handling techniques obtained by watching an e-learning tutorial.

Keywords:

Eye-tracking, e-Learning, Moving and Lifting Patients.

Introduction

It is estimated that about 40 % of all work injuries in the Danish healthcare system are caused by accidents related to moving and handling patients [1]. Interventions to reduce the risk of work injuries caused by the moving and handling of patients are often multi-dimensional. They generally include both the use of new patient-handling equipment, increased knowledge about safe moving and handling techniques and education and training in both use of equipment and techniques [2]. However, the principles and techniques taught on single courses are often not implemented in the workplace [3]. The limited effect of single courses calls for a more systematic and ongoing education regarding the use of patient-handling equipment and safe moving and handling techniques.

E-leaning in healthcare

E-learning is being used increasingly in healthcare as a flexible method to provide continuous education to healthcare professionals. Educational topics include practical skills such as safe moving and handling techniques. Over the last decades, the effect of e-learning in healthcare has been studied extensively. While some studies have failed to find evidence for the effect of e-learning compared to traditional education, others have found that e-learning had a positive effect on student knowledge and skills [4-6]. Despite discrepancies in the literature regarding the effects, there are several other advantages related to this type of education, such as flexibility and the ability to provide education for larger groups [6].

Only a few studies have assessed the effect of e-learning tutorials teaching safe moving and handling techniques. Harrington et al. concluded that there was a significant increase on the knowledge, and practices of nurses after completing an e-learning tutorial [7]. Similar results were found by Hayden et al., while Anderson et al. found that e-learning tutorials gave a better common interdisciplinary understanding and knowledge on safe moving and handling techniques [8,9].

The burden on the healthcare systems caused by the growing population of elders and dependents entail an increasing demand on healthcare professionals to possess the most recent knowledge in all areas. This demand can be difficult to accommodate by traditional education only [10]. Despite the uncertainty on the effect of e-learning tutorials in healthcare, it may be reasonable to expect that the use of e-learning in healthcare will continue to increase. Applying e-learning, it is possible to provide health care professionals with the most recent and relevant knowledge, despite time or location.

Eye tracking

Eye tracking has for the last decades been increasingly used to evaluate human-computer interaction [11]. Eye tracking is a measurement technology that can be used to obtain knowledge about where a person is looking on a motive of interest, usually a screen, for how long and in what sequences [11]. These measurements can be used to identify and analyze a person's visual attention or cognitive load while reading or searching the Internet [12]. In eye tracking, eye movements are typically processed and analyzed in fixations

or saccades. A fixation is the period where the eye is relatively still, in order to capture the visual stimuli. A saccade is the series of quick movements that occur between two fixations. During a saccade the visual acuity is suppressed, meaning that visual stimuli are perceived during the fixations [12].

Most eye tracking studies deploy the eye-mind hypothesis published in 1980 by Just and Carpenter. According to this hypothesis, longer fixations are associated with a longer cognitive processing load [13]. Later studies have shown other results, but the general understanding is that the duration of a fixation can be used as an indicator of the degree of visual attention and of the complexity of the cognitive process, e.g. in learning [14,15]. However, the literature is scarce on eye tracking for evaluation of the obtained knowledge after using an e-learning tutorial. One study by Soh et al. tested student knowledge before and after completing an e-learning tutorial. In this study, the students both had a significant increase of 45% in the mean number of fixations and a decrease of 49% in the mean time to first fixation after using an e-learning tutorial [16]. The use of eye tracking for evaluating an e-learning tutorial on practical skills, such as safe moving and handling techniques, has not been found in the literature. Therefore, it was relevant to investigate the usability of eye tracking as a measurement technology for evaluating e-learning tutorials related to practical skills.

The aim of this study was to explore whether eye tracking in combination with the qualitative methods observation and interviews can be applied to assess practical skills obtained after watching e-learning videos regarding safe moving and handling techniques. Combining quantitative and qualitative methods using method triangulation, we could obtain a complementary objective and subjective knowledge on the correlation between a person´s visual attention, practical performance and experience when using an e-learning tutorial regarding safe moving and handling techniques.

Materials and Methods

Participants

Four participants accepted to participate in the study. The mean age was 29.5 (±4.5) years. All four participants were students from the master program in Clinical Science and Technology at Aalborg University, Denmark. Their bachelor degree varied, two of the participants were bachelors in nursing, one was an occupational therapist, and the last was a radiographer. The experience level among the participants varied, but none of the participants had attended any courses or had worked related to safe moving and handling techniques within the last two years.

Equipment

For this study we used a Tobii X2-30 eye tracker from Tobii Technology (Tobii AB, Stockholm, Sweden). The Tobii X2-30 has a sample rate of 30 Hz with an operating distance from eye tracker to participant on 40–90 cm (15.7–33.5″) and a gaze angle up to 36° [17]. The eye tracker was mounted at the bottom of a Dell Latitude E6540 laptop with a screen size of 15,6".

E-learning tutorial

The Danish website, forflyt.dk, is administered by a non-profit organization (The Community of Work Environment for Welfare and Public Administration). This website contains an e-learning tutorial on safe moving and handling of patients. From this e-learning tutorial we chose 4 different videos, that demonstrated safe moving and handling techniques used in four different situations. The four situations were repositioning up in bed, hoisting from the floor, going from laying in a bed to standing upright and hoisting from bed to a wheelchair [18]

Each video contained both small video clips and still pictures with written text. Both the video clips and still pictures contained spoken text.

From the four videos we predefined 12 Areas of Interests (AOI). The predefined AOIs were chosen from a theoretical point of view and were all related to the demonstration of the correct use of different patient-handling equipment

Still pictures illustrating the 12 AOIs are shown in figure 1 and 2.

Figure 1-The AOI´s from video 1 and video 2

Figure 2- The AOI´s from video 3 and video 4

Study Design and Procedure

The procedure consisted of four sessions.

Before the beginning of the first session, a calibration procedure was carried out for each participant. During this calibration procedure the Tobii eye tracker measures the characteristics of the participants eyes and uses them together with an internal, physiological 3D eye model to calculate the gaze data. This individual model contains information on light refraction, shapes and reflection properties on different parts of the eye e.g. the placement of fovea, cornea etc. The measurements are stored in the Tobii Pro Studio Software, and therefore this procedure was only conducted prior to the participants watching of the first video [19].

In each session, the participants first watched a video demonstrating the use of patient-handling equipment and safe moving and handling technique in relation to a specific

situation. While watching the videos, the participants' eye movements were measured using the Tobii X2-30 eye tracker. After watching the video, the participant performed the same safe moving and handling techniques as seen in the video.

After performing the safe moving and handling techniques, participants were interviewed on four questions relating to the video just seen.

Data Analysis

The data from the eye tracking recordings, the observations and interviews were all analyzed separately.

Data from eye tracking recordings

Data from eye tracking recordings were processed in the software program Tobii Pro Studio version 3.4.5. In this program, we drew the 12 AOIs, for all the 16 recordings (4 recordings for each of the four participants). Since there was a minor displacement between the video frames for the 16 recordings, the AOI was drawn one frame and one recording at a time. This way we could do the necessary relocations of the AOI according to the action in the underlying video. Only one AOI were activated at a time.
We also used Tobii Pro Studio to calculate the metrics for the individual and total fixation duration, and the number of fixations for the 12 AOIs from each participant.
Before calculating the metrics, we applied the Tobii I-VT filter. Amongst the I-VT filter presets, we chose to alter the following. Noise reduction was set to apply MovingMedian and a window size of 3 samples. Eye selection was set to Average to calculate data from both eyes when possible and on only one eye when both eyes could not be used. The Discard short fixations were set to discard fixations shorter than 80 ms.

The calculated metrics on the AOIs were then classified into seen or not seen using a threshold of 300 ms. If the participant's fixation on an AOI was longer than 300 ms, the fixation was classified as seen and given the value 1, and a 0 if the fixation were less than 300 ms.

Observations

The 16 recorded videos from the participants' practical performances were viewed and then classified as correct or incorrect performance. Correct performance was given the value 1 and an incorrect performance was given the value 0.

Interviews

The 16 interviews were processed as described by Kvale. The interviews were first transcribed and then coded into central themes. The central themes were then condensed into descriptive statements [20].

Combining data

The processed data from both the eye tracking and the observations for the AOIs of each video was combined in a matrix for each video. We then compared the matrix for each video and the interview data from the same video to qualify the matrix results using the participants' subjective statements.

Results

Video 1

The participants expressed that they obtained most knowledge from seeing the video, but also expressed that they found it difficult to see how specific techniques were performed. As shown in table 1, there was no correlation between seeing an AOI and performing it correctly, as none of the participants performed the moving and handling technique correctly.

Table 1- The results from eye tracking data and findings from observations, from video 1 (AOI no. 1-3)

Video 1	See		
Perform		Yes	No
	Yes	0	0
	No	6	6

Video 2

The results from this video indicated a minor correlation between seeing an AOI and performing it correctly, as shown in table 2 Data from the interviews showed that the participants felt that they obtained most knowledge from what they saw, but also what they heard. As reported on video 1, the participants were not always able to see clearly what was demonstrated. Also, the design of the video (mainly stops and repetitions) apparently made it difficult for the participants to remember what they had seen.

Table 2-The results from eye tracking data and findings from observations, from video 2 (AOI no. 4 – 7)

Video 2	See		
Perform		Yes	No
	Yes	7	5
	No	2	2

Video 3

Again, the participants reported that it was difficult to see the details, and that the lack of spoken text and design of the video made it difficult to remember. According to the results shown in table 3, only one AOI had been seen, while six AOIs had been performed correctly, indicating that there was no positive correlation between seeing an AOI and performing the technique correctly.

Table 3- The results from eye tracking data and findings from observations, from video 3 (AOI no. 8 – 9)

Video 3	See		
Perform		Yes	No
	Yes	1	5
	No	0	2

Video 4

Results from video 4, as seen in table 4, showed clear relations between having seen an AOI and performing it correctly and between not seeing it and not performing it correctly. Central themes found in the interviews were that the participants thought that even though they obtained most knowledge from seeing the video, there were still elements that they found disturbing, especially that the video depicted a healthcare professional playing a demented woman instead of just demonstrating a more neutral situation.

Table 4-The results from eye tracking data and findings from observations, from video 4 (AOI no. 10 – 12)

Video 4	See		
Perform		Yes	No
	Yes	5	2
	No	4	1

Discussion

The ability to perform safe moving and handling techniques requires practical skills and thus motor skills. Key factors when learning motor skills are the complexity, difficulty and recognizability of the specific task [21,22]. Two of the AOIs in video 1 were related to the use of a piece of slide sheet, which was folded using a special technique. Our results showed that even though two of the participants saw the AOIs, none of the participants performed the technique correctly. This could indicate that the task demonstrated in the AOIs was too difficult and complex for the participants to perform after seeing it only once. This was also reported by the participants in the following interview. When scoring the performance from the observation data, a correct performance required the participant to use the same technique as demonstrated in the video and not just any technique leading to the correct function of the slide sheet. If the performances were scored correct when the participant got the slide sheet to function in first attempt, the distribution of the matrix would have been different, as there would have been several of the participants who would have performed the task correctly.

The impact of complexity, difficulty and recognizability of tasks when learning motor skills may also influenced the results in video 2. As seen in Table 2 there was a minor tendency for a positive correlation between the number of AOIs that had been seen and a correct performance. Several of the AOIs for this video demonstrated the use of movements such as rolling, pushing and dragging. These movements contain motor skills that are probably more recognizable and intuitive for the participants, which could have had, an effect on their ability to transfer their knowledge from what they saw on the video to the performance [21,22]. This indicates that the AOIs were adequate in terms of complexity of the task demonstrated.

Movements relating to what is known as the natural pattern of movement may have influenced the results in video 3.

Even though the participants expressed that they obtained most knowledge from what they saw, only one of the AOIs for that video was seen by a participant. Despite only one AOI being seen, the AOIs had been performed correctly in six out of eight cases. Both AOIs in this video were related to getting the patient's legs over the side of the bed, which is a common and intuitive movement for most people [21,22].

Another factor that may have influenced the results and findings could be the design of the videos in terms of camera angles and the use of spoken text.
As mentioned before, only one AOI in video 3 had been seen by a participant. In this video the two AOIs demonstrated an action related to the legs of the patient, which in the video were seen in on the left side corner. Studies show that people tend to gaze at faces and upper bodies and that they tend to gaze in the same area on a screen when seeing a scene [23,24]. It could be that the participants, when watching the video, have had their gaze on the patient's and assistant's faces and upper bodies and that they kept their attention in that area for this reason. Despite that there has only been one registered fixation above the 300 ms. threshold, it does not necessarily mean that the actions of the AOI were not perceived by the participants. Viewers are namely able to perceive and extract the gist of a scene from a brief 40-100 ms exposure [25]. It is possible that the participants, even by very short fixations got enough visual information to be able to perform the safe moving and handling technique correctly.

All videos involved a speaker telling what is being done in the video. The use of spoken text seems to cause increased visual attention [15,26]. In video 1, AOI no. 2 demonstrates an assistant moving a pillow, which is also told but without explaining why it is moved. All though this was seen by two participants, none of the four participants performed this AOI correctly. This could indicate that the verbal cue by the spoken text might have been to simple, not pointing out the relevance of why the pillow is moved. Also, the sequence is short, which could have caused participants to overlook it. Another example of an AOI not been performed by most of the participants is AOI no. 5 in video 2. This AOI demonstrates the how to apply a sling on a patient. The demonstration is not accompanied by spoken text but is quite long and takes 69 seconds. It was seen by three of the participants, but only performed correctly by one. This could indicate that despite the sequence being long, the lack of spoken text as a cue for guiding the visual attention could influence the participants' ability to perform it correctly. This could be supported by the results from video 4 where AOI no. 11 also demonstrates the application of a sling, but in this video, the demonstration is accompanied by spoken texts. All four of the participants saw the AOI and three performed it correctly, indicating that the spoken text had a positive influence on participant performance.

The results reported here suggest that even though the participants expressed obtaining most learning form what they saw, there were still a several factors, affecting their practical performance. Amongst the most important factors were difficulty in seeing how patient-handling equipment were used and the lack of spoken text to explain the correct usage. Also, the design of the video according to still pictures, repetitions i.e. were experienced by the participants

as disturbing and making it difficult to remember the content of the video. As for the data from the eye tracking recordings and observations there were no clear correlation in the four videos, which was probably due to several factors, such as our choice of AOI according to the difficulty and complexity of the demonstrated technique, the difference in the design of the videos and our criteria for scoring the observations.

Conclusion

Including eye tracking in method triangulation has a potential for providing knowledge on the correlation between what a person sees, performs and experience, i.e. what people unconsciously focus on vs. the intended learning outcomes in the e-learning tutorial. This can be used to optimize usability tests and e-learning tutorials, in which the main purpose is to facilitate learning involving physical performance.

Limitations and further studies

Based on this study's strengths and limitations and the lack of similar studies, further research using a corresponding approach could provide more data to examine if it is possible to detect a more clear and detailed pattern, involving duration of fixations and the physical performance, to make stronger conclusions to validate eye tracking as a tool in assessing e-learning tutorials. Further studies could also include RCT studies, where pre- and posttest can determine whether there are differences in fixation times for respectively a control group that receives traditional education and an experimental group that receives education by an e-learning tutorial.

Acknowledgement

References

[1] Kines P, 5000 arbejdsulykker blandt SOSU´er – Nationale Forskningscenter for Arbejdsmiljø Available at:http://www.arbejdsmiljoforskning.dk/da/projekter/sosu/laes-mere-om-arbejdsmiljoeet-i-aeldreplejen/fysisk-arbejdsmiljoe/5000-arbejdsulykker-blandt-sosuer Assessed 2018 May 3

[2] Lunde PH, Kranker S, Plambech L, Bøgedal G, Nylander P, *Forflytningskundskab: Aktivering, hjælp og træning ved forflytning*. København: Gads Forlag; 2010

[3] Clemes SA, Haslam CO, Haslam RA. What constitutes effective manual handling training? A systematic review. Occupational Medicine. 2010;60(2):101-107

[4] Chumley-Jones HS, Dobbie A, Alford CL. Web-based Learning: Sound Educational Method or Hype? A Review of the Evaluation Literature. Academic Medicine. 2002;77(10):86-93

[5] Lahti M, Hätönen H, Välimäki M. Impact of e-learning on nurses´ and student nurse's knowledge, skills, and satisfaction: A systematic review and meta-analysis. International Journal of Nursing Students. 2014;51(1):136-149

[6] Vaona A, Banzi R, Kwag K, Rigon G, Cereda D, Pecoraro V, Tramacere I, Moja L. E-Learning for health professionals. Cochrane Database or Systematic Reviews. 2018.

[7] Harrington SS, Bonnie LW, Teaching ergonomics to nursing facility managers using computer-based instruction. Journal for Nurses in Staff Development. 2006;22(5):260-268

[8] Hayden C. Online Learning of Safe Patient Transfers in Occupational Therapy Education. The Open Journal of Occupational Therapy. 2013;1(2): 1-20

[9] Anderson MP, Safe moving and handling of patients: an interprofessional approach. Nursing Standard. 28, 46, 37-41.

[10] Elearning Industry. Top 5Growing Industries Using eLearning – eLearning Industry [internet]. 2018 Available at: https://elearningindustry.com/growing-industries-using-elearning-top-5. Assessed 2018 May 19

[11] Bojko A, *Eye tracking the user experience: a practical guide to research*. Rosenfeld Media; 2013

[12] Tobii Technology. Tobii Eye Tracking An introduction to eye tracking and Tobii eye Trackers. [internet] 2010. Available at: www.Tobii.com Assessed 2018 March 1

[13] Just Ma, Carpenter PA. A Theory of reading: Form eye fixations to comprehension. Psychological Review. 1980; 87(4)329-354.

[14] Liu W, Yu M, Fan Z, Xu J, Tian Y. Visual attention based evaluation for multiple-choice tests in e-learning applications. Proceedings - Frontiers in Education Conference, FIE. 2017. p. 1-6

[15] Yang F-Y, Tsai M-J, Chiou G-L, Lee SW-Y, Chang C-C, Chen L-L. Instructional Suggestions Supporting Science Learning in Digital Environments Based on Review of Eye Tracking Studies. Journal of Educational Technology & Society 2018;21(2):28-45

[16] Soh BP, Reed WM, Poulos A, Brennan PC. E-Tutorial improves students´ ability to detect lesions. Radiologic Technology. 2013;85(1):17-26

[17] Tobii Technology. Tobii Pro X2-30 screen-based eye tracker – product description [internet]. 2015 Available at: https://www.tobiipro.com/product-listing/tobii-pro-x2-30/#Specifications. Assessed 2018 May 26

[18] BrancheFællesskabArbejdsmiljø, Velfærd of Offentlig Administration. E-læringsmodul. [internet]. 2018. Available at: https://www.forflyt.dk/materialer_og_vaerktoejer/e_laeringsmodul. Assessed 2018 April 16

[19] Tobii AB. Tobii Studio User´s Manual Version 3.4.5. 2017 [internet]. Available at: https://www.tobiipro.com/siteassets/tobii.../user-manuals/tobii-pro-studio-user-manual.pdf. Assessed 2018 March 1

[20] Kvale S, Brinkmann S. *Doing Interview*. Hans Reitzel: 2009

[21] Voelcker-Rehauge C. Motor-skills learning in older adults – a review of studies on age related differences. European Review of Aging and Physical Activity. 2008;5(1):5-6

[22] Newell KM. Motor Skill acquisition. Annual Review of Psychology. 1991;42(1):213-37

[23] Goldstein RB, Woods RL, Peli E. Where people look when watching movies: do all the viewers look at the same place? Computers in biology and medicine. 2007;37(7):957-64

[24] Nielsen J, Pernice K. *Eyetracking web usability*. New Riders; 2010 p. 437

[25] Rayner K, Smith TJ, Malcolm GL, Henderson JM. Eye movements and visual encoding during scene perception. Psychological science. 2009;20(1):6-10

[26] Jamet E. An Eye-tracking study of cuing effects in multimedia learning. Computers in Human Behavior. 2014; 32 pp. 47-53

A Systematic Review of Cluster Detection Mechanisms in Syndromic Surveillance: Towards Developing a Framework of Cluster Detection Mechanism for EDMON System

Prosper Kandabongee Yeng[a], Ashenafi Zebene Woldaregay[a], Terje Solvoll[b], Gunnar Hartvigsen[a]

[a] *Department of Computer Science, University of Tromsø – The Arctic University of Norway, Tromsø, Norway*
[b] *Norwegian Centre for E-health Research, University Hospital of North Norway, Tromsø, Norway*

Abstract

Time lag in detecting disease outbreaks remains a threat to global health security. Currently, our research team is working towards a system called EDMON, which uses blood glucose level and other supporting parameters from people with type 1 diabetes, as indicator variables for outbreak detection. Therefore, this paper aims to pinpoint the state of the art cluster detection mechanism towards developing an efficient framework to be used in EDMON and other similar syndromic surveillance systems. Various challenges such as user mobility, privacy and confidentiality, geographical location estimation and other factors have been considered. To this end, we conducted a systematic review exploring different online scholarly databases. Considering peer reviewed journals and articles, literatures search was conducted between January and March 2018. Relevant literatures were identified using the title, keywords, and abstracts as a preliminary filter with the inclusion criteria and a full text review were done for literatures that were found to be relevant. A total of 28 articles were included in the study. The result indicates that various clustering and aberration detection algorithms have been developed and tested up to the task. In this regard, privacy preserving policies and high computational power requirement were found challenging since it restrict usage of specific locations for syndromic surveillance.

Keywords:

Syndromic Surveillance, Spatiotemporal Clustering, Smart Phone, Aberration Detection.

Introduction

Late detection of disease outbreak has been a threat to global health security for quite a long time, which cost the world many lives, resources, fear and panic. Case fatality rate (CFR) of pandemic diseases is still in the ascendance. The most recent being Ebola Virus Disease (EVD) in Liberia, West Africa. Apart from global fear and panic, EVD registered over 11000 deaths with national case fatality rate of about 70% and local economic losses of $3-4 billion [1, 2]. Traditional surveillance systems are mostly passive and rely on laboratory confirmations to detect disease outbreak. This

has been enhanced to syndromic surveillance systems [3] which largely depends on visible signs and symptoms with data sources including emergency department records [4], school absenteeism, work absenteeism, disease reporting systems and over-the-counter medication sales [5, 6]. Nevertheless, the existing syndromic surveillance systems could not detect the disease outbreak early enough and their data sources, and process excludes the incubation phase of the infection [6]but efforts are being made to bridging the gap [6-9].

Recently, the availability of the internet and ubiquity of systems such as smart phones, tablets, smart watches, laptops and other systems have created greater opportunity for the advancement of diabetes management technologies and this generates big data [10]. In the right mix of cluster detections, big data from self-management of diabetes, internet availability and the prevailing pervasiveness of devices, it is feasible and efficient to detect infectious disease outbreak as early as the incubation stage by using the vulnerability of diabetes patients as a sensor [7]. Detection of disease outbreak at the incubation stage is important for reducing morbidity and mortality through early prevention and control [11-14]. Therefore, the general objective is to conduct a systematic review to determine the state-of-the-art clustering detection method, design and evaluation strategies. Associated challenges such as user mobility, privacy and confidentiality along with estimation of geographical location towards the development of a cluster detection approach for EDMON and other similar syndromic surveillance systems would be pinpointed.

Clustering Approach and Outbreak Detection

Generally, outbreak of diseases could be presented in cluster form either in space, time, or both [15, 16]. Clustering methods in disease outbreak detection helps in the identification of environmental factors and spreading patterns linked with certain diseases [10]. Clustering approach could be roughly categorized as temporal, spatial and spatio-temporal. Spatial clustering uses multi-dimensional vectors with longitudinal and latitudinal coordinates. There are variety of such algorithm such as density-based spatial clustering of applications with noise (DBSCAN) [15-17]. Temporal clustering deals

with data points associated with time [18, 19]. It includes various algorithms such as cumulative summation (CUMSUM) and what is strange about recent event (WSARE) [20-22]. Spatiotemporal clustering occurs when there is the involvement of time and spatial dimension [15-17]. There are variety of strategies including different distance functions [23, 24], importing time to the spatial data, transform spatiotemporal data to the new objects, progressive clustering and spatiotemporal pattern discovery [15, 17]. Aberration detection is mainly performed through thresholding mechanisms including various forms such as number of standard deviation set from the mean (z-score), generalized likelihood ratio (GLR), recurrence interval (RI) and confidence intervals (CI) [25, 26].

Materials and Methods

A literature search was conducted between January 2018 and March 2018 through Google Scholar, Science Direct, Pub-Med, IEEE, ACM Digital Library and Scopus. Different key words such as "Spatiotemporal Clustering" /" Syndromic Surveillance", "Syndromic Surveillance"/ "trajectory Clustering", "Real Time"/"Syndromic surveillance"/"Clustering Mechanism", "Cell Phone"/"Syndromic surveillance"/"Clustering", "Mobile Phone"/"Syndromic surveillance"/Clustering", "Smart Phone"/"Syndromic surveillance"/"Clustering", and "Aberration Detection"/"Syndromic Surveillance"/"Clustering" were used. For a better searching strategy, key words were combined using Boolean functions such as 'AND' 'NOT, and 'OR'. Peer reviewed journals and articles were considered. The inclusions and exclusions criteria were developed based on the objective of the study and through rigorous discussions among the authors. Guided with the inclusion and exclusion criteria, basic filtering was done by skimming through the titles, abstracts and keywords to retrieve records which seemed relevant. Duplicates were removed and articles, which seems relevant, based on the inclusion and exclusion criteria, were fully read and judged. Other relevant articles were also retrieved using the reference list of accepted literatures. Preferred Reporting Items for Systematic Reviews and Meta-Analysis (PRISMA) flow diagram was used to record the article selection and screening [27].

Inclusion and Exclusion Criteria

For an article to be included in the review, the study should be a practically implemented system with cluster detection mechanisms. Practically implemented algorithms were being sorted for because the results of the study were intended to be used for development of a framework and practical implementation of syndromic surveillance system in EDMON and such similar systems. The study did not have sufficient resources to explore into theoretical and unimplemented algorithms for practical implementations hence the need to skew to practically implemented algorithms. The study was also limited to English language as it does not have the required resources needed to evaluate and accommodate participants who do not speak or write English[28]. The publication type included journal articles, conference abstracts and

presentations. There were no time restrictions. Any other article outside the above stated scope were excluded

Data Collection and Categorization

The data collection and categorization were developed based on the objective, literature reviews and authors discussions. The categories have been defined exclusively to assess, analyzed and evaluate the study as follows:

Table 13-Data categories and their Definitions

Category	Definition
Clustering and Aberration Detection Algorithm	This category defines the kind of clustering and Aberration detection algorithm which the study has used and implemented.
Type of Clustering Algorithm	This category defines the type of algorithm. The type of algorithms includes spatial, temporal and spatiotemporal algorithms.
Threshold	This category defines the type of threshold used to generate alarms and alerts in the study.
Clustering Category	The clustering algorithms has been categorized [15]. This dimension tags the specified clustering algorithm used to their respective category.
Design method	This category indicates the design method such as prototype, participatory or joint application development, Agile or waterfall model, that has been used in implementing the system.
Evaluation criteria	In this category, the evaluation criteria used in evaluating the algorithms has been specified.
Performance metrics	This category specifies the performance metrics such as sensitivity, specificity, positive predictive value etc., which was used in the evaluation of the algorithms.
Type of Location	Different type of locations are being used in clustering. These include geolocation, postcodes, counties and many others. This category specifies the exact type of location which was used in the system.
Source of Location	The source of location is defined as the location where the type of location information was obtained from.
Nature of Location	The nature of the location is defining the state of the location as static or dynamic nature.
Visualization tool used	This category also records the type of visualization tool used in the implementation of the visualization aspect of the system.

Display Report	This category records the type of visual displays (graphs, maps, time series etc.) which were implemented by the various systems in the study.
Design layout	This category records the stages and processes used in the architectural design of the syndromic surveillance system. For example, a layout may consist of data acquisition, clustering and aberration detection and visualization [25]. While other design layout could include privacy preserving mechanisms, machine learning techniques in processing the data and other layers [29, 30].

Literature Evaluation and Analysis

Eligible literatures were assessed, analyzed and evaluated, based on the above defined categories. Analysis was performed on each of the categories (Clustering and Aberration Detection Algorithm, Type of clustering, Threshold, Cluster Category, Location Type, Location Source, Nature of Location Source, Design Method, Evaluation, Visualization Tool, Display Report and Design Layout) to evaluate the state of the art approaches. Percentages of the attributes of the categories were calculated based on the total number of counts (n) of each type of the attribute. It is better to take a note that some studies might use multiple categories, therefore, the number of counts of these categories could exceed the total number of articles of these systems presented in the study.

Results

Relevant Literatures

Through searching in the various online databases, a total of 5,936 records were found. Reading of titles, abstracts, keywords and guided by the inclusion and exclusion criteria, led to an initial exclusion of 5,793 literatures and further removal of duplicates in the record resulted in 125 literatures, which were fully read and judged. After full text reading, a total of 28 articles were included in the study and analysis as shown in figure 3.

Table 1. Type of Clustering Algorithms.

Type of Algorithms	Usage Count	%
Spatial	16	32
Spatiotemporal	19	38
Temporal	15	30

Table 3: Type of Threshold Detection Mechanisms

Threshold	Usage #	%
Confidence Interval (CI)	1	4
Generalized Likelihood Ratio (GLR)	5	19
Incidence Ration (IR)	1	4
Recurrence Interval (RI)	10	38
Z-Score	9	35
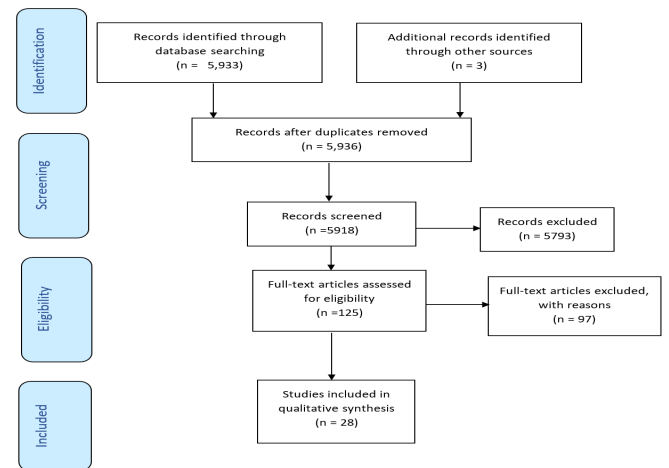
Figure 3: Flowchart of the review process.

Literature Evaluation and analysis

As described earlier, the literatures were assessed, analyzed and evaluated based on the above defined categories. The following section will describe the findings.

I. Types of Clustering Algorithms

Among the three types, namely spatial, temporal and spatio-temporal clustering algorithm, spatio-temporal algorithm is found to be the most preferred approach followed by spatial and temporal algorithm respectively as shown in the table 1.

II. Clustering and Aberration Detection Algorithms

There are a variety of clustering and aberration detection algorithms implemented in the reviewed literatures, where space-time permutations scan statistics is widely adopted followed by cumulative summation, space-time scan statistics and others as shown in the table 2.

Table 4: Categories of Clustering Algorithms.

Algorithm Category	Usage Count	%
Different Distance Function (DDF)	1	3
Importing Time to Spatiotemporal Data (ITTSD)	12	32

Spatiotemporal Pattern Discovery (STPD)	2	5
Threshold base Clustering (TBC)	23	60

Table 2- Clustering and Aberration Detection Algorithms.

Algorithm	Usage Count	%
Risk-Adjusted Support Vector Cluster-ing(RSVC)	1	2
Bayesian Spatial Scan Statistics (BSSS)	1	2
Cumulative Summation (CUMSUM)	8	16
DBSCAN	2	4
Exponentially Weighted Moving Average (EWMA)	1	2
Flexible Space-Time Scan Statistic(FSTSS)	1	
kernel Density	3	6
K-means clustering	1	2
K-NN with Haversian distance (KNearness)	1	2
Log-Linear Regression(LLR)	3	6
Moving Average(MA)	2	4
(Shewhart Chart (P Chart)	1	2
Pulsar Method(PM)	1	2
Recursive Least Square(RLS)	2	4
Risk Adjusted Nearest Neighbor Hierarchical Clustering (RNNH)	1	2
Small Area Regression and Testing (SMART),	1	2
Statistical Process Control(SPC)	2	4
Space Scan Statistics(SSS)	3	6
ST-DBSCAN	1	2
space-time permutation scan statistic(STPSS)	9	18
Space-time Scan Statistics(SSS)	5	10
What is Strange About Recent Event (WSARE)	1	2

III. Threshold Detection Mechanisms

Aberration detection is mainly performed using thresholding mechanisms and in this regard, there are various types of approaches implemented in the reviewed literatures. To this end, Recurrence Interval (RI) is the most widely adopted strategies followed by Z-score, GLR and others as shown in the table 3.

IV. Categories of Clustering Algorithms (CCA)

There are various categories of clustering algorithms, from which threshold-based clustering is the most widely adopted as shown in table 4.

V. Design, Evaluation Methods and Performance Metrics

The reviewed literatures have used various evaluation strategies, among which simulation with historical data stood out as the most widely adopted approach as shown in table 6. The performance metrics which were mostly used are sensitivity (44%) and specificity (36%) as shown in table 5. Prototype and participatory designed were used in the study. Out of 5 systems which disclosed their design methods, 4 of them used participatory approach.

Table 5- Performance Metrics

Performance Metric	Usage Count	%
Sensitivity	11	44
Specificity	9	36
Timeliness	2	8
Consistency	1	4
Correlation	1	4
Positive Predictive Value	1	4

Table 6- Evaluation Method

Evaluation Type	Usage Count	%
Simulation with Historical Data	12	80
Comparison with Known Outbreak	2	13
Power of Cluster Detection Test	1	7

VI. Location Type & Nature, and Source of Location

The literatures have used variety of location type, nature and source as shown in the table 7-9. In this regard, majority of the study used static location (79%) and the rest used dynamic location (21%). Moreover, the study exploited various address such as Geocode (50%), Zip Code (46%) and County (4%). Furthermore, various source of locations has been explored such as Patient Health Record (64%), Mobile Device (14%), TCP/IP (11%), County (4%), and School Address (4%).

Table 7- Location Type

Type of Location	Usage Count	%
Geocode	14	50
Zip Code	13	46
County	1	4

Table 8- Nature of Location Type

Nature of Location	Usage Count	%
Static	22	79
Dynamic	6	21

Table - Source of Location

Source of Location	Usage Count	%
Patient Health Record	18	64
TCP/IP	3	11
Mobile Device	4	14
County	1	4
School Address	1	4

VII.   Visualization Tools and Visual Displays

Clustering and aberration detection mechanisms in diseases outbreak needs to be backed up with excellent visualization tools and display to facilitate a quick response from the concerned bodies on the exact timing and place. In this regard, the reviewed literatures have adopted various kinds of tools, among which ArcGIS (24%), Google Map API (22%), Twi-Info (22%) as shown in table 10. Moreover, as to the displaying mechanisms, map (47%) is the most widely used followed by time series (27%), graphs (23%) as shown in table 11.

Table 10- Visualization Tools

Visualization Tool	Usage Count	%
ArcGIS	3	24
Google Map API	2	22
TwiInfo	2	22
OpenStreetMap	1	11
JFreeChart	1	11

Table 11- Visual Displays

Visual Display	Usage Count	%
Maps	14	47
Time Series	7	27
Graphs	8	23
Color Indicators	1	3

VIII.   Design layout

The design layout identified in the study have been abbreviated and defined as follows;

DCADAA: This layout consists of obtaining Data first. Then Clustering and Aberration detection are done, followed by generating Alarms to create   Alerts of aberrations [20]. DCAVAA: A visualizing module is built in addition, to processes defined in DCADAA [29]. DCTCAVAA: In addition to DCAVAA layer defined above, this layer has data cleaning and transformation features. DCFADAA: In addition to DCADAA, this layout does data filtering or categorizing the data into some defined groups either manually or by employing machine learning techniques. DPVCAAA: In addition to DCAVAA layout, this layout has privacy preserving mechanisms such as anonymization and pseudonymizing [31, 32]. RDPVCAAA: On top of DPVCAAA layout, there is an additional module which for real time data process[31] [29, 31]. TDCAVVAA: In addition to DCAVAA, this layout, tracks user's movement to obtain the data. This is followed with validating the data before Clustering and Aberration detection.[29, 30].

Table 12- Design Layout.

Abbreviation	Usage Count	%
DCADAA	12	55
DCAVAA	1	4.5
DCTCAVAA	3	14
DCFADAA	2	9
DPVCAAA	2	9
RDPVCAAA	1	4.5
TDCAVVAA	1	4.5

Discussion

The general objective is to use a systematic review to assess the state-of-the-art clustering algorithms and other features of systems, which can be used to develop an effective and efficient cluster detection mechanism in EDMON and other similar syndromic surveillance systems. A summary of the most

used approaches and categories are given in the table 13 below;

Table 13: Summary of the most used approaches

Category	Most Used
Clustering Algorithm	Space Time Permutation Scan Statistics
Type of Clustering	Spatiotemporal type
Threshold	Recurrence Interval
Algorithm Category	Threshold base Clustering
Design Method	Participatory Design
Evaluation Method	Simulation with historical data
Performance Metric	Sensitivity
Type of Location	Geocode
Source of Location	Patient Health Record
Nature of Location Source	Static
Visualization Tool Used	ArcGIS
Displayed Output	Maps
Layout	DCADAA

STPSS is one of the spatiotemporal algorithms which is used by most of the syndromic surveillance systems in detecting disease outbreak. Space and time of potential disease outbreak detection is a very efficient method since health management can plan for such potential outbreaks. Health management would know where and when to allocate resources to potential outbreak areas. Another reason of its high usage count could be that the algorithm does not require population at risk data to draw the expected baseline value. But it dwells on the detected cases to determine the expected count [14]. This approach provides significant trend of baseline data while avoiding inclusion of historical data that is irrelevant to the current period. STPSS unlike most of the algorithms does not draw its baseline data (expected cases) from inaccurate population at risk, a control group, or other data that provide information about the geographical and temporal distribution of the underlying population at risk. Such baseline data are inaccurate because there exist significant geographical variation in health-care utilization data due to differences in disease prevalence, health care access and consumer behavior [14]. Unlike spatiotemporal algorithm, spatial algorithms would only indicate where aberrations would occur. This makes planning difficult for health management since it will be difficult to know when to implement health interventions having known potential places for disease outbreak. Sometimes, spatial algorithms are implemented together with temporal algorithms [33]. This gives the surveillance system the spatiotemporal properties. The most used

thresholds for aberration detection in spatiotemporal algorithms was Recurrence Interval (RI). This could be as a result that the combination of RI and Monte Carlo Replication helps to easily determine and set specificity of the system [34]. The Monte Carlo simulation is a probability module which is often used with RI or GLR on a cluster to draw a threshold and to determine the likelihood occurrence of a cluster by chance within a specified period for which the analysis is repeated in a regular basis. For instance, in a daily analysis, if the Monte Carlo replication was set to 999 with statistically significant signal of p value < 0.001, the RI would be 1000 days since in disease surveillance the RI is the inverse of the p value. [34]. This implies that, for each 1000 day, the expectation of false alarms would be an average of one false signal per 1000 days or 2.7 years and the RI would be set to the number of days of the baseline data[35].

CUMSUM is a temporal algorithm which was mostly used together with special algorithms. Its ease easy and efficiency might have accounted for the high usage[36].About 60% of the algorithms were classified to be Threshold Based Category (TBS) [15]. This corresponded to relatively high usage of spatiotemporal algorithms. Most of these algorithms employed cylindrical risk regions to detect clusters. The radius formed the area of the map, while the height represented the time. The radius and time were varied to some upper bound thresholds. Participatory design was majorly used while simulation with historical data was mostly used to evaluate the clusters in most of the algorithms. Sensitivity and specificity were the most used performance metrics in the evaluation. This could be the case because users were possibly much interested in a system with reduced false alarms rate. In terms of location, geocodes of census track or hospitals and zip codes were mostly used as location points for the clustering algorithms. These records were mostly retrieved from patient health records. Dynamic nature of the sources of location were of low count. The low count could have been due to the undeveloped and difficulties associated in acquiring and processing dynamic nature of location source data for syndromic surveillance. Also, the stringent inclusion and exclusion criteria on practically implemented syndromic surveillance systems might have accounted for the low count of dynamic nature of location sources. Furthermore, privacy preserving polices and high computational time requirement prohibited the use of exact location of persons for syndromic surveillance. Exact locations such as house numbers and tracking of individuals were only used for group data at the zip code or county level. Information on the exact place of infection is also vital for early prevention and control of morbidity and mortality. But these limitations often hamper the accuracy of information on place of infection since the information collected often relates much to the place of notification which is usually far from place of infection [37, 38]. Also, systems which provided text space for users to indicate their location had some limitations. Users did not indicate proper locations or addresses so their locations could not be geocoded. This resulted in limited sample size [32, 39].

ArcGIS was mostly used to display graphs in this review. It is possible that maps were majorly displayed because it can be used to represent both spatial and spatiotemporal data.

This could have accounted for their high usage of 34% and 47% in their respective categories. In the system design layout category, most of the systems were interested in obtaining data from various sources first. Clustering and Aberration detection were done, followed by generating Alarms to create Alerts of aberrations. This was abbreviated to (DCADAA) for ease of data processing. Tracking for data, acquiring data in real time, privacy preserving mechanisms, filtering and data cleaning were some of the layout processes employed in few of the systems studied. The low rate of tracking persons for data sources could be due to legal, privacy and ethical reasons. Low count of filtering and data cleaning could be due to implementation challenges as machine learning algorithms and natural language processing tools are used for effectiveness. Privacy preserving mechanism is also very vital of which all the systems should have implemented [31]. But the low count rate could have been due to low enforcement of privacy preserving laws in data processing.

The Study Limitations

There is a limitation resulting from impact and study design[40].The study was specifically focused on practically implemented algorithms in relation to syndromic surveillance using clustering mechanisms. The inclusion and exclusion criteria were very specific and stringent on practically implemented syndromic surveillance systems. Therefore, there is the tendency of missing out some algorithms which were not practically implemented in syndromic surveillance systems. For instance, despite an exhaustive search in combination with the search keys, "Cell Phone", "mobile phone" and "Smart Phone", there were limited information regarding mobile phone base trajectories clustering used in syndromic surveillance.

Conclusion

The aim of this review was to derive the state-of-the-art clustering algorithm and its associated design and evaluation methods from practically implemented syndromic surveillance systems. The study revealed Space-Time Permutation Scan Statistics as the most implemented algorithm. The uniqueness and efficiency of STPSS is that its baseline or expected count is based on its detected cases within a defined geographical distance (cylinder radius) and area or temporal window (cylinder height). This approach provides significant trend of baseline data while avoiding inclusion of historical data that is irrelevant to the current period. This algorithm can be used in EDMON and other similar syndromic surveillance systems that are aiming towards implementing state-of-the-art cluster detection mechanism. Temporal and spatial algorithms can also be combined to achieve efficient space time result. This study has also provided wide data categorization, ranging from design of the system to the display of reports. Therefore, we foresee these results might foster the development of effective and efficient cluster detection mechanisms in EDMON and other similar syndromic surveillance systems.

References

[1] WHO. *Ebola Virus Disease*. 2017 June 2017 [cited 2018 20/01/2018]; Available from: http://www.who.int/mediacentre/factsheets/fs103/en/.

[2] Daulaire, N.M., *Global Health Security*. 2018.

[3] Hope, K., et al., *Syndromic surveillance: is it a useful tool for local outbreak detection?*, in *J Epidemiol Community Health*. 2006. p. 374-5.

[4] Choi, J., et al., *Web-based infectious disease surveillance systems and public health perspectives: a systematic review*. BMC Public Health, 2016. **16**(1): p. 1238

[5] Nie, S., et al., *Real-Time Monitoring of School Absenteeism to Enhance Disease Surveillance: A Pilot Study of a Mobile Electronic Reporting System*, in *JMIR Mhealth Uhealth*. 2014.

[6] Woldaregay, A.Z., et al. *EDMON-A Wireless Communication Platform for a Real-Time Infectious Disease Outbreak De-tection System Using Self-Recorded Data from People with Type 1 Diabetes*. in *Proceedings from The 15th Scandinavian Conference on Health Informatics 2017 Kristiansand, Norway, August 29–30, 2017*. 2018. Linköping University Electronic Press.

[7] Heffernan, R., et al., *Syndromic surveillance in public health practice, New York City*. Emerg Infect Dis, 2004. **10**(5): p. 858-64,15200820,

[8] Jacquez, G., *Spatial Clustering and Autocorrelation in Health Events | SpringerLink*. 2018

[9] Woldaregay, A., et al., *An Early Infectious Disease Outbreak Detection Mechanism Based on Self-Recorded Data from People with Diabetes*. Studies in health technology and informatics, 2017. **245**: p. 619-623

[10] Wang, H. and U.o.S.C.-. Columbia, *Pattern Extraction From Spatial Data - Statistical and Modeling Approches*. 2014, University of South Carolina.

[11] MedicineNet, *Modeling Infectious Diseases in Humans and Animals*. 2017.

[12] Study.com. *Progress of Disease: Infection to Recovery - Video & Lesson Transcript | Study.com*. 2018; Available from: http://study.com/academy/lesson/progress-of-disease-infection-to-recovery.html.

[13] Marshall, J.B., et al., *Prospective Spatio-Temporal Surveillance Methods for the Detection of Disease Clusters*. 2009

[14] Martin Kulldorff, R.H., Jessica Hartman, Renato Assunção, Farzad Mostashari, *A Space–Time Permutation Scan Statistic for Disease Outbreak Detection*. 2005

[15] Fanaee-T, H., *Spatio-Temporal Clustering Methods Classification (PDF Download Available)*, in *Doctoral Symposium on Informatics Engineering*. 2012.

[16] P.N. Tan, Vipin Kumar, and M. Steinbach, *Cluster Analysis: Basic Concepts and Algorithms*. 2005

[17] Birant, D. and A. Kut, *ST-DBSCAN: An algorithm for clustering spatial–temporal data*. Data & Knowledge Engineering, 2007. **60**(1): p. 208-221

[18] Hutwagner, L., et al., *Comparing Aberration Detection Methods with Simulated Data*, in *Emerg Infect Dis*. 2005. p. 314-6.

[19] Chan, T.C., Y.C. Teng, and J.S. Hwang, *Detection of influenza-like illness aberrations by directly monitoring Pearson residuals of fitted negative binomial regression models*, in *BMC Public Health*. 2015.

[20] Kleinman, K.P., et al., *A model-adjusted space-time scan statistic with an application to syndromic surveillance*. Epidemiol Infect, 2005. **133**(3): p. 409-19,15962547,2870264.

[21] Kulldorff, M., *A spatial scan statistic*. http://dx.doi.org/10.1080/03610929708831995, 2007

[22] Chen, D., et al., *Spatial and temporal aberration detection methods for disease outbreaks in syndromic surveillance systems*. http://dx.doi.org/10.1080/19475683.2011.625979, 2011

[23] Khokhar, S. and A.A. Nilsson. *Introduction to Mobile Trajectory Based Services: A New Direction in Mobile Location Based Services*. in *Wireless Algorithms, Systems, and Applications*. 2009. Berlin, Heidelberg: Springer Berlin Heidelberg.

[24] Jeung†, H., et al., *Discovery of Convoys in Trajectory Databases*. 2008

[25] Sharip, A., *Preliminary Analysis of SaTScan's Effectiveness to Detect Known Disease Outbreaks Using Emergency Department Syndromic Data in Los Angeles County*. 2006.

[26] Kajita, E., et al., *Harnessing Syndromic Surveillance Emergency Department Data to Monitor Health Impacts During the 2015 Special Olympics World Games*. Public Health Rep, 2017. **132**(1_suppl): p. 99s-105s,28692391,PMC5676508.

[27] PRISMA. *PRISMA*. 2018; Available from: http://www.prisma-statement.org/.

[28] Omicsonline. *Inclusion and Exclusion Criteria and Rationale*. 2018; Available from: https://www.omicsonline.org/articles-images/2157-7595-5-183-t001.html.

[29] Ali, M.A., et al., *ID-Viewer: a visual analytics architecture for infectious diseases surveillance and response management in Pakistan*. Public Health, 2016. **134**: p. 72-85,26880489,

[30] Groeneveld, G.H., et al., *ICARES: a real-time automated detection tool for clusters of infectious diseases in the Netherlands*. BMC Infect Dis, 2017. **17**(1): p. 201,28279150,PMC5345172.

[31] GDPR, E. *EU GDPR Information Portal*. 2018; Available from: http://eugdpr.org/eugdpr.org.html.

[32] Yan, W., et al., *ISS--an electronic syndromic surveillance system for infectious disease in rural China*. PLoS One, 2013. **8**(4): p. e62749,23626853,PMC3633833.

[33] Khanita Duangchaemkarn Varin, C.P., Wiwatanadate, *Symptom-based data preprocessing for the detection of disease outbreak - IEEE Conference Publication*, in *2017 39th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*. 2017: Seogwipo. p. 2614-2617.

[34] Takahashi, K., et al., *A flexibly shaped space-time scan statistic for disease outbreak detection and monitoring*. International Journal of Health Geographics, 2008. **7**(1): p. 14

[35] Yih, W.K., et al., *Evaluating real-time syndromic surveillance signals from ambulatory care data in four states*. Public Health Rep, 2010. **125**(1): p. 111-20,20402203,PMC2789823.

[36] Hutwagner, L., et al., *The bioterrorism preparedness and response Early Aberration Reporting System (EARS)*. J Urban Health, 2003. **80**(2 Suppl 1): p. i89-96,12791783,PMC3456557.

[37] Cesario, M., et al., *Time-based Geographical Mapping of Communicable Diseases - IEEE Conference Publication*. 2012

[38] Qi, F. and F. Du, *Tracking and visualization of space-time activities for a micro-scale flu transmission study*. International Journal of Health Geographics, 2013. **12**(1): p. 6

[39] Nicholas Thapen, et al., *DEFENDER: Detecting and Forecasting Epidemics Using Novel Data-Analytics for Enhanced Response*. 2016

[40] Edanz Group Japan K.K., *Writing Point: How to Write About Your Study Limitations Without Limiting Your Impact | Edanz Editing*. 2015

Address for correspondence:

Prosper Kandabongee Yeng,
MSc (Information and Network Security)
Department of Computer Science
University of Tromsø - The Arctic University of Norway
Realfagbygget Hansine Hansens vei 54 Breivika
Tromsø, 9019
Norway
Phone: 47 96992748
Email: prosper.yeng@gmail.com/pye000@post.uit.no

Developing a Bayesian Network as a Decision Support System for Evaluating Patients with Diabetes Mellitus Admitted to The Intensive Care Unit – a Proof of Concept

Rune Sejer Jakobsen, Ole Hejlesen, Mads Nibe Stausholm, Simon Lebech Cichosz

Department of Health Science and Technology, Aalborg University, Aalborg, Denmark

Abstract

Evidence is increasing about an unsatisfying performance from the existing non-disease-specific scoring systems in the intensive care unit (ICU). Evidence is furthermore increasing about differences in the mortality rate between diabetics and non-diabetics dependent on the level of blood glucose (BG), but few scoring systems include these variables in the assessment of the patients. 142,404 ICU admissions were included from the eICU database in the development of an unsupervised trained Bayesian Network (BN). The BN suggested that abnormalities in the level of BG should be associated with differences in the mortality rate between diabetics and non-diabetics. The BN showed promising predictive ability with an AUC on 0.86 for predicting death (sensitivity: 75.06, specificity: 78.40 %). 48.43 % of the length of stays (LOS) were correctly predicted. The results were slightly below the results from the APACHE IV scoring system but showed great ability of risk stratification. The BN showed a potential for predicting the patient outcome and might enable an improved method for risk stratifying the patients admitted to the ICU.

Keywords:

Intensive Care Unit, APACHE IV, Mortality, Diabetes Mellitus, Blood Glucose.

Introduction

The ICU is a preventive and therapeutic care setting that take care of patients who suffers from a very vulnerable and potential reversible failure of the organ systems that is more severe and beyond a treatable limit from a regular bed department [1] [2]. The immune system of the patients at the ICU is often at a worsened state [1]. The ICU includes observation, diagnosis, treatment and caretaking of the patient [2].

In the last century has the ICU been under a rapidly development, including a development of multiple scoring systems [3]. Most of the different scoring systems includes a score – an integer – and a probability model, which predict the mortality rate for the patient [3].

Most of the existing scoring systems are developed to evaluate general disease severity for a mixed patient group in the ICU [4]. Some studies suggest, that the accuracy of the models is sufficient when evaluating a mixed group of critical ill patient admitted to the ICU [5] [6]. On the other hand, is the evidence also increasing about the models performs insufficient related to disease specific complications, such as Cardiac Arrest, H1N1 (influenza A), Acute Respiratory Distress Syndrome and patients requiring respiratory support [7] [8] [9] [10] [11]. This might support the discussion of whether the old models developed for evaluating general disease severity is suited for evaluating specific disease severity in the ICU [4].

A disease-specific patient group, meeting an increased interest in the ICU, is the patients suffering from Diabetes Mellitus (DM), where abnormalities in the level of BG are reported to have different impact on the mortality rate dependent on whether the patient suffers from DM [12] [13]. DM is the seventh most common cause of death in USA and is furthermore related to other comorbidities, such as retinopathy, neuropathy, renal failure, coagulative- and cardiovascular complications [14]. Although, the actual impact from DM, and the underlying mechanisms, on both the severity of the condition and the mortality rate not fully understood [15]. It might therefore be essential develop a model that enables an analysis of correlations between changes in multiple variables for the patients with DM. The disease-specific scoring system for diabetics should additionally be able to analyze the joint probability of DM and the level of BG in the ICU to improve the understanding of how patients suffering from DM proceeds in the ICU compared to non-diabetics.

The objective of this paper was therefore to develop a new model for decision support and risk stratification of patients suffering from DM admitted to the ICU. The model should enable an assessment of the impact from a joint probability of DM and the level of BG and furthermore seek to improve the accuracy of the predictions regarding mortality rate and LOS, compared to the APACHE IV scoring system.

Materials and Methods

The data acquired was collected from 2014-2015 and contains data about critical ill patients admitted to the ICU in

USA. Data was collected in the eICU Collaborative Database. The eICU Collaborative Database included data from > 200,000 patient admissions. The mean age of all the patients was 61.97 years, 45,90 % was female and 54.10 % was male.

Data screening

142,404 ICU admissions were registered with information of whether the patients suffered from DM and an actual mortality rate and an actual LOS. Admissions with missing information regarding DM, mortality rate and LOS were excluded. 108,899 admissions were registered without DM and 33,505 admissions registered with diabetes. There was no discrimination between Type 1 Diabetes Mellitus (T1DM) and Type 2 Diabetes Mellitus (T2DM) in the eICU database. 4.93 % of the patients suffering from DM died in the hospital and 6.17 % of the patients without DM died in the hospital.

Bayesian Network

A BN was used as it enabled an assessment of correlations in data, including analyzing the joint probability of two or more variables in the ICU. BN's used the Bayesian rule to update the probability of each node when evidence e was presented. The Bayesian rule stated:

$$P(x|e) = \frac{P(e|x)P(x)}{P(e)} \quad (1)$$

Which was that the probability of having the disease given symptoms, $P(x|e)$, could be calculated if the probability of having the symptoms, $P(e|x)$, the probability of having the disease, $P(x)$, and the probability of having the symptoms, $P(e)$, were known.

When evidence was presented in the network, would the probabilities be updated dependent on how the connections in the network was modelled. This concept was called probabilistic inference. The probabilistic inference was used for reasoning in the network. The Bayesian rule can be rewritten as [16]:

$$P(x|e) \propto P(e^c|x) \times P(x|e^P) =$$
$$\prod_{j=1}^{|C|} P(e_{C_j}|P_{mn}) \left[\sum_{all\,P_{mn}} P(x|P_{mn}) \sum_{j=1}^{|C|} P(P_i|e_{p_i}) \right] \quad (2)$$

This indicated, that the probability of having the disease given symptoms, $P(x|e)$, had a proportional dependency on the children nodes, $P(e^C|x)$, and the parent nodes, $P(x|e^P)$. [16] Information could not flow between children nodes in a BN, which instead was written as the product of all the likelihoods, $\prod_{j=1}^{|C|} P(e_{C_j}|P_{mn})$. C_J represented the j 'th child note, e_{C_J} represented the value of the probabilities of cardinality it states. P_{mn} represented the value of the state n on the parent node P_m. All priors were summarized and multiplied by their condition probability, $\sum_{all\,P_{mn}} P(x|P_{mn}) \sum_{j=1}^{|C|} P(P_i|e_{p_i})$. The children nodes and the parent nodes were normalized to get exactly $P(x|e)$. [16]

The Bayesian rule and the probabilistic inference were used to let information pass though the network and predict how evidence affect the probabilistic outcome throughout the entire network [16]. This means that a quantification of the condition, a risk stratification and a predicted score for mortality rate and LOS could be calculated dependent on entered evidence e in the BN.

Structure learning

The development of the graphical structure of the BN was performed in Hugin Expert Developer (https://www.hugin.com/). Structure learning was performed unsupervised using the Greedy Search-and-score algorithm provided by Hugin Expert. A maximum of 13 parents were chosen in the Greedy Search-and-score algorithm for the BN and Bayesian Information Criterion (BIC) was chosen as the penalized likelihood criteria. BIC penalizes the complexity of the model and rewards the model if it fits the data well. [17]

Parameter learning

The distributions of data-points between the states in each node – the parameters – were learned by using the Expectation-Maximization (EM) algorithm. The EM algorithm handled missing data very well, making it suited for handling real-life data. The EM-algorithm was furthermore considered as the natural choice of algorithm for parameter learning if the structure learning was performed by the Greedy Search-and-Score algorithm [18]. The EM-algorithm counts the parameter for the BN.

Validation

The holdout method was used for validation, which separated the dataset into a training set and a test set [19]. The data was distributed as 90 % for training and 10 % for test. This method ensures that the model was less likely to be overfitted.

The accuracy of the predictions from the BN was compared with the predictions from the APACHE IV scoring system. Two APACHE IV models were developed for the test: one for mortality and one for LOS. The two APACHE IV models were illustrated in Figure 1.
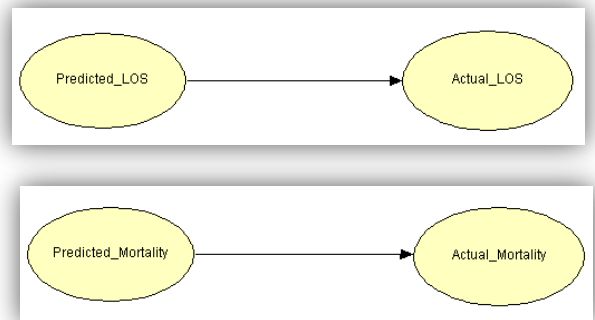

Figure 1 – Presented the two models for testing the accuracy of the predictions from APACHE IV

The same unique patientunitstayID was used for both testing the BN and the APACHE IV scoring system. 128,164 admissions were used for training and 14,240 admissions were used for testing.

Results

Three tests were performed for the BN: (1) One for testing the impact of a joint probability of DM and the level of BG on the mortality rate, (2) one for testing the performance of predicting the mortality rate and (3) One for testing the performance of predicting the LOS.

The impact from DM and level of BG on the mortality rate

Figure 2 illustrates the impact on the mortality rate from a joint probability of the level of BG and respectively diabetics and non-diabetics.
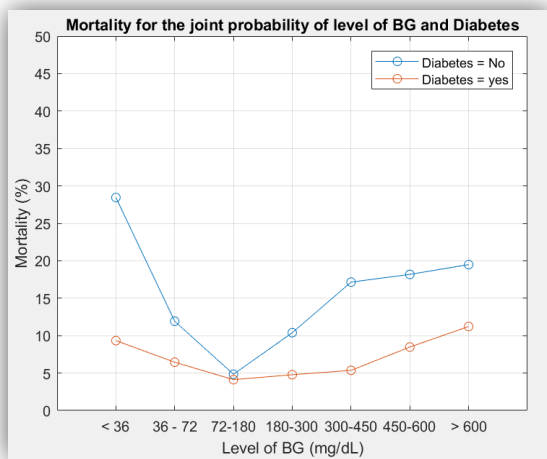

Figure 2 – presented the correlation between mortality rate and the joint probability of the level of BG and diabetes. 72 – 180 mg/dL was associated with a normal level of BG

Performance test for predicting mortality rate

A performance test was also carried out to evaluate the accuracy of the predictions from the BN regarding the mortality rate for both diabetics and non-diabetics, which were also compared with a performance test for the APACHE IV scoring system. The performance test consisted of a ROC-curve with an associated AUC, and a result table which illustrated the relationship from different cut-offs between sensitivity and specificity. The ROC-curve for predicting mortality for both diabetics and non-diabetics were a test of the accuracy of predicting the [mortality: dead]. The ROC-curve was presented in Figure 3:
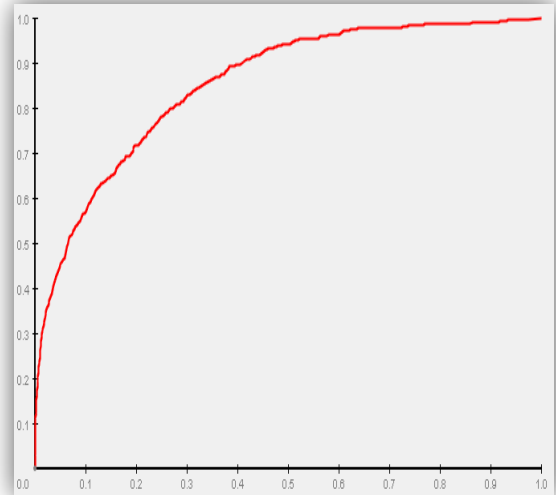

Figure 3 – Presented the ROC-curve of [mortality: dead] with an AUC on 0.85561. The ROC-curve illustrates sensitivity (true positive rate) vs. 1 – specificity (false positive rate)

The ROC-curve had an AUC on 0.86 for the BN developed in this project. In comparison did the test of APACHE IV show an AUC on 0.87 for [Mortality: Dead], which was tested following Figure 1.

Different cut-offs could be chosen for sensitivity and specificity for the BN developed in this project. A moderate sensitivity and a moderate specificity also calculated because a high sensitivity caused a low specificity and a high specificity caused a low sensitivity. The relationship between sensitivity and specificity were presented in a result table, which was presented in Table 1:

Table 1 – Presented the relationship between sensitivity and specificity in terms of number of true positive (No. TP), number of false positive (No. FP), number of true negative (No. TN) and number of false negative (No. FN).

	Sensitivity	Specificity	No. TP	No. FP	No. TN	No. FN
High Sensitivity	100.00 %	0.12 %	12416	823	1	0
High Specificity	2.71 %	100.00 %	364	0	824	13052
Moderate Specificity	75.06 %	78.40 %	10070	178	646	3346

The APACHE IV scoring system had a moderate sensitivity on 75.57 % and a moderate specificity on 81.07 %.

Performance test for predicting LOS

The performance test of predicting LOS consisted of a confusion matrix, where predictions was from a max belief and a

calculation of the percentage of correct predictions. The confusion matrix for LOS was illustrated in Table 4:

Table 2 – Presented the confusion matrix of LOS

< 1. (actual)	1. – 2. (actual)	> 2. (actual)	
808	608	425	< 1. (predicted)
1640	1754	1442	1. – 2. (predicted)
1215	2029	4318	> 2. (predicted)

48.43 % were correctly predicted from a test of max belief. In comparison did the test of APACHE IV show that 52.54 % was correctly predicted from a test of max belief, which was tested following Figure 1.

Discussion

An AUC > 0.80 was associated with good discrimination capability [5], meaning the accuracy of predicting the mortality rate was considered satisfying for the BN even though the accuracy of the predictions from the BN was slightly below the predictions from the APACHE IV scoring system. The accuracy of predicting the LOS was considered satisfying, because the test was performed as a test of max belief, whereas the actual prediction from the BN was a distribution of likelihoods. The accuracy from the test of max belief of the prediction of LOS was slightly below the predictions from the APACHE IV scoring system.

The literature suggested that the non-diabetics was more vulnerable towards hyperglycemia and hypoglycemia [12, 13] which was also suggested by the BN. The increased mortality for non-diabetics with abnormalities in the level of BG might be explained by the diabetics are more accustomed towards both hyperglycemia and hypoglycemia [15].

It was possible that the mortality would be different for T1DM and T2DM, where e.g. the body mass index of especially T2DM could have an impact on the mortality rate [15]. Hba1c could also be a part of the assessment of the patient, which could explain occurrence of abnormalities in the level of BG prior to ICU admission or even cases of unknown DM [14]. This might affect the degree of how accustomed the patient was to experiencing abnormalities in the level of BG. Intensive insulin treatment could also be within the scope of assessing diabetics in the ICU, as the mortality rate and LOS might be dependent on the intensive insulin treatment [20] and because the intensive insulin treatment was strongly associated with normal DM caretaking [21, 22].

The Greedy Search-and-score algorithm was used for structure learning, but other algorithms exist for unsupervised structure learning. The PC-approach makes edges between all nodes with no conditional independence in the BN [18], meaning a very complex BN will be generated with very high number of CPT's, which would be practically impossible to compile for a BN with many nodes and states. This could be solved by inspecting the model and deleting some of the edges, which might hold the potential of ensuring all edges associated with domain knowledge were included in the model. It would be interesting for a future study to evaluate whether the PC-approach could improve the accuracy of the predictions, as an alternative structure of the BN could be produced [18]. The Rebane-Pearl might hold the potential for improving the transparency of the model, where edges were directed away from a root where one node could only have one parent (except for the root-node) [18]. The Rebane-Pearl algorithm was not used for unsupervised structure learning, as one patient could have more diagnoses and one variable for APS was associated with other APS variables for the same organ system, meaning one node was expected to have multipleparents.

The Greedy Search-and-Score algorithm was considered suited for weighting the accuracy and transparency, where an inspection of the generated structure for a disease-specific model for DM showed promising edges. The Greedy-search-and-Score did also show a promising accuracy of the predictions, meaning the algorithm was considered as a promising approach for performing unsupervised structure learning.

Conclusion

The output of this paper was a BN developed through unsupervised structure learning (by the Greedy Search-and-score algorithm) and parameter learning (by the EM algorithm). The BN was able to quantify the severity of the condition, predict the outcome for the patient and potentially improve the risk stratification of the patient admitted to the ICU. The BN could possibly be a means for decision support for the clinician when evaluating patients suffering from DM admitted to the ICU.

The results showed high difference in the likelihood for mortality rate for the joint probability of DM and the level of BG.

Both the BN and the APACHE IV scoring system showed high accuracy in predicting the mortality rate and relatively poor accuracy of predicting the LOS, though predicting LOS for both the BN and the APACHE IV scoring system were better than just guessing. The accuracy of the predictions of mortality rate and LOS from the BN was slightly below the accuracy of the predictions from the APACHE IV scoring system. The BN was associated with an improved and alternative method for risk stratification where the BN was able to present changes in the likelihood from correlations between evidence in multiple variables for both diabetics and non-diabetics in the ICU.

References

[1] Last JM. A Dictionary of Public Health. Oxford University Press. 2007.

[2] Dansk Intensiv Database. Årsrapport. 2016.

[3] Rapsang AG, and Shyam DC. Scoring systems in the intensive care unit: A Compendium. 2014.

[4] Vincent JL, and Moreno R. Clinical review: Scoring systems in the critically ill. 2010.

[5] Zimmerman JE, Kramer AA, McNair DS, and Malila FM
Acute Physiology and Chronic Health Evaluation (APACHE) IV: Hospital mortality assessment for today's critically illpatients. 2006.

[6] Naqvi IH, Mahmood K, Ziaullaha S, Kashif SM, and Sharif A. Better prognostic marker in ICU - APACHE II, SOFA or SAPS II!. 2016.

[7] Enfield K, Miller R, Rice T, Thompson BT, and Truwit J.
Limited Utility of SOFA and APACHE II Prediction Models for ICU Triage in Pandemic Influenza. 2011.

[8] Saleh A, Ahmed M, and Mansoura AA.Comparison of the Mortality Prediction of Different ICU Scoring Systems (APACHE II and III, SAPS II, and SOFA) in Acute Respiratory Distress Syndrome Patients. 2016.

[9] Shahpori R, Stelfox HT, Doig CJ, and Boiteau PJE. Sequential Organ Failure Assessment in H1N1 pandemic planning. 2011.

[10] Bisbal M, Jouve E, Papazian L, Bourmont SD, Perrin G, Eon B, and Gainnier M. Effectiveness of SAPS III to predict hospital mortality for post-cardiac arrest patients. 2014.

[11] Rojek-Jarmuła A, Hombach R, and Krzych ŁJ. APACHE II score cannot predict successful weaning from prolonged mechanical ventilation. 2017.

[12] Tayek CJ, and Tayek JA. Diabetes patients and non-diabetic patients intensive care unit and hospital mortality risks associated with sepsis. 2012.

[13] Cichosz SL, and Schaarup C. Hyperglycemia as a Predictor for Adverse Outcome in ICU Patients With and Without Diabetes. 2017.

[14] Carpenter DL, Gregg SR, Xu K, Buchman TG and C. M. Coopersmith. Prevalence and Impact of Unknown Diabetes in the ICU. 2015.

[15] Siegelaar SE, Devries JH, and Hoekstra JB. Patients with diabetes in the intensive care unit; not served by treatment, yet protected?. 2010.

[16] Duda RO, Hart PE, and Stork DG. Pattern Classification. 2001.

[17] Cho SJ, and Boeck PD. A Note on N in Bayesian Information Criterion for Item Response Models. 2018.

[18] Hugin Expert. Hugin Expert – Manual. 2017.

[19] Schneider J. Cross Validation. 1997.

[20] Barski L, Kezerle L, Zeller L, Zektser M, and Jotkowitz A. New approaches to the use of insulin in patients with diabetic ketoacidosis. 2013.

[21] Almdal T, Kjeldsen HC, Vestergaard H, and Sachs C. Sundhed.dk - Type 1 diabetes. 2016.

[22] Almdal T, Kjeldsen HC, Sanbæk A, Vestergaard H, and Sachs C. Sundhed.dk - Type 2 diabetes. 2015.

Predicting Preventable Hospitalizations among Elderly Recipients of Home Care: a Study Protocol

Mads N. Stausholm, Pernille H. Secher, Simon L. Cichosz, Ole K. Hejlesen

Department of Health Science and Technology, Aalborg University, Aalborg, Denmark

Abstract

Ageing population and traditional consequences of ageing is expected to cause an increased number of recipients of community care services and socio-economic burden. To accommodate these changes, The Danish Health Authority has recommended several tools to help community personnel detect early signs of disease in the home care setting. These tools are used solely for detecting current deviations from the habitual health status, and no data analysis is performed in order to predict upcoming deviations. This paper describes a study protocol to investigate the potential of developing a data driven decision support model to predict unplanned, preventable hospitalizations. Machine learning techniques, such as logistic regression, will be applied on data from three various sources in order to predict which recipients of home care services are at risk of an unplanned, preventable hospitalization. If successful, the proposed model may facilitate earlier prediction and actions towards deviations from the individual citizen's habitual health status and thereby increase the chance of prevention of hospitalization and functional decline.

Keywords:

Health Services for the Aged, Decision Support Techniques, Forecasting, Hospitalization.

Introduction

The number of recipients of home care services and the socio-economic burden hereof is expected to increase due to the combination of population ageing and traditional consequences of ageing [1]. Rising life expectancy and declining fertility rates are among the main contributors of population ageing. Traditional consequences of normal ageing are decreases in health status, physical function, and quality of life [1], [2], [3]. Such decreases can lead to acute events that may result in unplanned hospitalizations. Hospitalization is associated with long-lasting or even permanent declines in health status and physical function [4], [5]. The aforementioned changes can affect the citizens' ability to conduct simple everyday tasks, such as cleaning their home or managing personal hygiene [1], and thereby cause a need for additional home care services. In Denmark, the home care services are primarily managed by the 98 municipalities, each taking care of its own residents. All 98 municipalities have agreed on a common definition of a preventable hospitalization, which is a hospitalization with one of the following primary diagnoses [6]:

- Nutritional anaemia
- Bone fracture
- Pneumonia
- Acute bronchitis
- Chronic lower respiratory disease
- Pressure ulcer
- Gastroenteritis
- Cystitis

On average, a Danish municipality will have some sort of home care interaction with 22% of its citizens aged 65 or above and 49% of those aged 80 or above [7]. According to [8], 2.4% of the entire Danish population either lives at an elderly care centre or receive home care services. These are responsible of 11.2% of the national health care expenses and the socio-economic burden hereof is expected to rise. The number of citizens aged 65 or above is expected to increase from 1.1 million (2017-level) to 1.6 million by year 2060 [9], and so is the number of care-needing citizens.

During the past few years the focus of detecting early signs of disease exacerbations and preventing unplanned hospitalizations in elderly has increased [10]. Several studies have investigated the risk factors of hospitalizations and prediction hereof and found several biomarkers, malnutrition, high and low BMI, comorbidities, renal insuffiency, polypharmacy, and hospitalization within the preceding year to be predictors of hospitalizations [11], [12], [13], [14].

In 2013, The Danish Health Authority published a report highlighting this problem and provided guidelines to help the community personnel assess and register the citizen's health information, and act upon any abnormalities detected [10]. Specific focus points were development of tools and instructions to further educate community personnel and help them

detect early signs of disease exacerbations and react properly thereto [10]. One of the tools provided to the community

While the aforementioned studies have identified risk factors and predictors of hospitalizations, most of these are not easily obtained in the current home care settings. In the municipality of Aalborg, the community personnel perform routine-
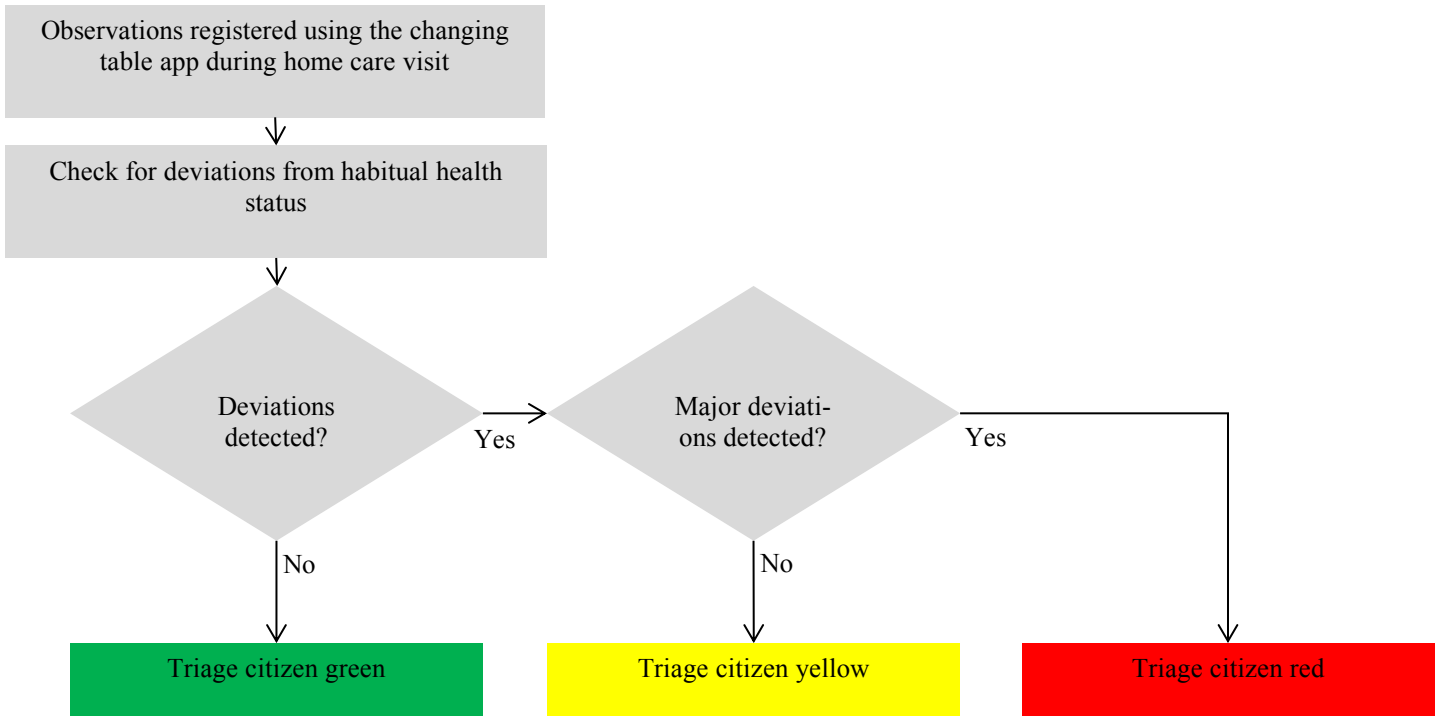

Figure 1 – Overview of the workflow when community personnel observe and register information in the changing table app and how the citizen is triaged based on presence deviations from habitual health status

personnel is the triaged changing table (Danish: Ændringsskema), which is a check-list tool used to detect changes in the citizen's habitual state by answering questions within five subgroups of health status:

- Mental and social state
- Behaviour in the home
- Activities of daily living
- Consumption of food and beverages
- Physical complaints

Figure 1 provides an overview of the workflow using the triaged changing table. The registrations are compared to the citizen's habitual health status and decision support are provided to the community personnel to help deciding if and which further actions are needed. Using these tools results in a much more detailed documentation of the five subgroups of health status of the individual citizen and provides useful information for a decision support system [10]. Based on the report from The Danish Health Authority, Aalborg Municipality implemented a digital version of the triaged changing table in order to minimize the number of preventable hospitalizations. In 2014, Aalborg Municipality provided home care for 7,657 citizens aged 65 or above, which in total experienced 834 preventable hospitalizations [6], [15], [16].

ly registrations during each home care visit. These registrations are used as input in a triage-model, which is designed to detect deviations from habitual health status and provide decision support for further actions. Implementation of the digital triaged changing table provides a systematic approach to identify deviations from the habitual health status and decision support regarding required actions. These tools are used solely for detecting current deviations from the habitual health status, and no data analysis is performed in order to predict upcoming deviations.

The research objective of this study is to investigate whether it is possible to use machine learning techniques to predict unplanned, preventable hospitalizations in elderly recipients of home care services based on demographic, municipal, and regional information combined with routinely collected community personnel registrations.

Materials and Methods

Study Design

This study will be conducted as a retrospective cohort study as historical data is used for investigating the potential in predicting preventable hospitalizations in elderly recipients of home care services. The study will be conducted within the municipality of Aalborg in the North Denmark Region.

Data from the period of 2016-2017 from at least 600 subjects is planned to be included in the study. The results of this study will be published in an international peer-reviewed journal.

Inclusion and Exclusion Criteria

All subjects must have received home care services in Aalborg Municipality at least once a week during 2016 and 2017 and reached the age of 65 as per 1st of January 2016. Furthermore, subjects must have experienced at least one preventable hospitalization during the period of interest. Terminal subjects and subjects with severe mental disorders will be excluded from the study.

Data

The data for this study is planned to be obtained from at least 600 recipients of home care services in the municipality of Aalborg. Data is extracted from three various sources: 1) The electronic home care record (EHCR), which holds both demographic and municipal information, such as past and current home care services, 2) the triaged changing table database where all registrations and triage state changes are stored, and 3) the Danish National Patient Register (DNPR), which holds information of e.g. hospitalizations. Table 1 shows which information variables that will be extracted from each of the three data sources. The data will be extracted as three different data sets and subsequently pooled based on subjects' social security numbers.

Table 1 - Overview of the extracted information from the three data sources

Data source	Information variables
Electronic home care record	Age, sex, gender, civil status, comorbidities, medical prescriptions, past and current home care services including start and stop dates, and cost of home care services
Triaged changing table database	Habitual health status, registrations from each home care visit, and information concerning deviations from the habitual health status
Danish National Patient Register	Previous hospitalizations including primary diagnosis, treatment, length of stay, and cost. Contacts with or services received by other primary care units

Classification

Pre-processing

As the subject is either at risk of an unplanned, preventable hospitalization the problem is considered a binary classification problem. The dataset will consist of demographic, municipal, and regional data as well as data from the triaged changing table database. Continuous municipal, regional, and triaged changing table data registrations will be divided into event and control periods:

- Event periods: Data from a fixed period, e.g. two weeks, leading up to a preventable hospitalization

- Control periods: Data which is not within an event period or immediately following hospital discharge

Control periods will be segmented using a sliding window approach with a window size of similar length as the event periods. Data analysis will reveal the appropriate window size. The number of control periods is expected to be much larger than the one for event periods.

Feature Selection

As shown in Table 1, the feature set will consist of nominal, ordinal, and continuous features from the EHCR, the triaged changing table database, and the Danish National Patient Register. The data will be divided according to the event and control period definitions.

The feature selection will be conducted using MATLAB R2017b by performing a sequence of backwards feature elimination and forward feature selection on the entire feature set. Features will be selected based on the minimization of the misclassification rate and the resulting feature set will be passed into the classification model.

Classification Model

The requirements of the classification model are robustness towards highly imbalanced data (event periods vs. control periods) and handling of various data types, such as nominal, ordinal, and continuous. A logistic regression model complies with the aforementioned requirements and will be the classification model of focus in this study. The logistic regression model returns the probability of the subject being at risk of experiencing a preventable hospitalization or not, which corresponds to the probability of the data sample being from an event or control period, respectively. This is illustrated in Figure 2, which also provides an overview of the model.
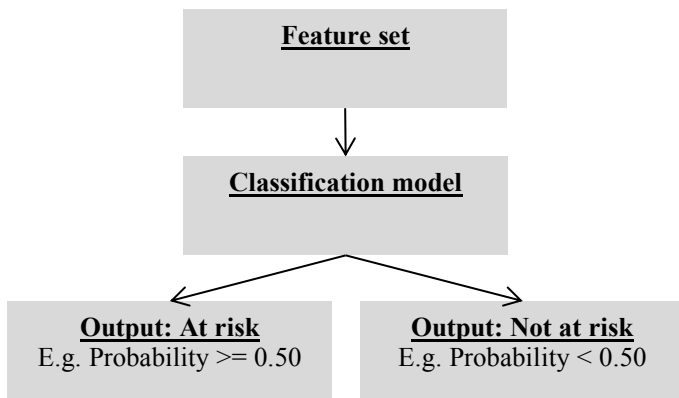
| Feature set |
| Classification model |
| Output: At risk
E.g. Probability >= 0.50 | Output: Not at risk
E.g. Probability < 0.50 |

Figure 2 – Overview of the classification model, which takes a feature set as input and provides a probability of an event as output

Cross validation will be performed in order to account for overfitting and the model performance will be evaluated using the averaged classification accuracy and standard deviation along with a confusion matrix and a receiver operating characteristic curve.

Implications

Reducing the number of unplanned, preventable hospitalizations in recipients of home care services is a national priority in Denmark and is considered a method to accommodate the consequences of population ageing. Currently several tools have been developed for this purpose but to our knowledge no initiatives involves the application of advanced data analytics and data science techniques. By applying such techniques on data that currently is registered continuously it may be possible to predict unplanned events at an earlier stage compared to the current practice. If successful, the proposed model may facilitate earlier actions towards deviations from the individual citizen's habitual health status and thereby increase the chance of preventing hospitalization and functional decline. This may lead to substantial societal gains why the study complies with the EU general data protection regulation.

References

[1] Ellenbecker CH, Samia L, Cushman MJ, and Alster K. Patient Safety and Quality in Home Health Care. In: Clancy CM, ed. Patient Safety and Quality: An Evidence-Based Handbook for Nurses: Agency for Healthcare Research and Quality 2008; pp. 301-340.

[2] Moen K, Ormstad H, Wang-Hansen MS, and Brovold T. Physical function of elderly patients with multimorbidity upon acute hospital admission versus 3 weeks post-discharge. Disability and Rehabilitation 2017: 8: 1-8.

[3] Sager M, Franke T, Inouye S, Landefeld C, Morgen T, Rudberg M, Sebens H, and Winograd C. Functional outcomes of acute medical illness and hospitalization in older persons. Archives of internal medicine 1996: 156(6); 645–652.

[4] Hoogerduijn JG, Schuurmans MJ, Duijnstee MSH, de Rooij SE, and Grypdonck MFH. A systematic review of predictors and screening instruments to identify older hospitalized patients at risk for functional decline. Journal of Clinical Nursing 2007: 16(1); 46-57.

[5] Buurman B, and de Rooij SE. Functional decline in older persons after acute hospitalization. Nederlands tijdschrift voor geneeskunde 2015: 159; A8185.

[6] KL - Local Government Denmark. Hvor god er kommunens hjemmepleje til at forebygge indlæggelser? – se nøgletallene for din kommune. KL - Local Government Denmark. Copenhagen, Denmark, 2016.

[7] The Danish Health Authority. Ældres sundhed – forebyggelse of sundhedsfremme i kommunerne. The Danish Health Authority. Copenhagen, Denmark, 2007.

[8] KL - Local Government Denmark, Danish Regions, Ministry of Finance, Ministry for Economic Affairs and the Interior, and Ministry of Health. Effektiv kommunal forebyggelse - med fokus på forebyggelse af indlæggelser og genindlæggelser. Ministry of Health. Copenhagen, Denmark, 2013.

[9] Statistics Denmark. FRDK117: population projections 2017 for the country by ancestry, sex and age. Statistics Denmark. Copenhagen, Denmark, 2017.

[10] The Danish Health Authority. Tidlig opsporing af sygdomstegn hos ældre medicinske patienter: Værktøjer til hverdagsobservationer. The Danish Health Authority. Copenhagen, Denmark, 2013.

[11] Umegaki H, Asai A, Kanda S, Maeda K, Shimojima T, Nomura H, and Kuzuya M. Risk factors for the discontinuation of home medical care among low-functioning older patients. Journal of Nutrition, Health and Aging 2016: 20; 453–457.

[12] McCusker J, Kakuma R, and Abrahamowicz M. Predictors of functional decline in hospitalized elderly patients: a systematic review. The journals of gerontology. Series A, Biological sciences and medical sciences 2002: 57(9); M569–M577.

[13] Stuck AE, Walthert JM, Nikolaus T, Büla CJ, Hohmann C, and Beck JC. Risk factors for functional status decline in community-living elderly people: a systematic literature review. Social Science & Medicine 1999: 48; 445–469.

[14] Rönneikkö JK, Mäkelä M, Jämsen ER, Huhtala H, Finne-Soveri H, Noro A, and Valvanne JN. Predictors for Unplanned Hospitalization of New Home Care Clients. Journal of the American Geriatrics Society 2016: 65(2); 407-414.

[15] Statistics Denmark. AED06: recipients referral to home care, free choice, by region, type of benefits, hours per week, age and sex. Statistics Denmark. Copenhagen, Denmark, 2017.

[16] Statistics Denmark. AED07: recipients referral to home care, nursing homes/nursing dwellings) by region, age and sex. Statistics Denmark. Copenhagen, Denmark, 2017.

Address for correspondence

Mads N. Stausholm, Department of Health Science and Technology, Aalborg University, Aalborg, Denmark;

E-mail: mns@hst.aau.dk.

Detection of Postprandial Hyperglycemia in Type 1 Diabetes Mellitus Patients – Initial Assessment of Current Recommendations versus Alternatives

Mia Birkholm Lausten, Ole Hejlesen, Mette Dencker Johansen

Department of Health Science and Technology, Faculty of Medicine, Aalborg University, Aalborg, Denmark

Abstract

To minimize the risk of microvascular and macrovascular complications, it is important to focus on detecting postprandial hyperglycemia, but it is only recommended to measure blood glucose (BG) 60 – 120 minutes after a meal if the preprandial BG is normal and HbA1c is increased. Thus the aim of this study was to investigate whether an alternative measurement interval of 120 – 180 minutes after a meal would outperform the recommended interval. Continuous glucose monitoring data, spot capillary blood glucose and carbohydrate intake times were collected from 23 type 1 Diabetes Mellitus patients. BG levels within 180 minutes after a meal were investigated, and the time the patients spent in postprandial hyperglycemia (BG >10 mmol/L) was converted to percentage for each interval and compared to all meals together, and to breakfast, lunch and dinner, respectively. No significant differences were found between the recommended interval and the alternative interval.

Keywords:

Postprandial Hyperglycemia, Recommendations, Type 1 Diabetes Mellitus, Continuous Glucose Monitoring, Blood Glucose Self-monitoring.

Introduction

Diabetes Mellitus (DM) is a huge healthcare problem worldwide as 425 million individuals in the age of 20 -79 years were affected in 2017 [1]. 5 - 10% of the patients have type 1 Diabetes Mellitus (T1DM) [2,3], where an auto-immune attack on the insulin secreting pancreatic β-cells leads to absolute absence of insulin [1,2,4]. T1DM can be effectively controlled and glycemic control can be achieved by a proper management of the blood glucose (BG) levels with daily injections of insulin [2]. Achieving good glycemic control is of great importance to minimize the risk of both microvascular and macrovascular long term complications in T1DM patients [2,5-7].

A clear association exists between glycated hemoglobin (HbA1c), which is average glycemic control over 2-3 months [1,8,9], and the risk of long term complications [10-13]. In addition, limiting the time spent in postprandial hyperglycemia is found to be an important contributor to improvement of HbA1c and thereby also to reduce the risk of long term complications [6,7,9,13-15]. Postprandial hyperglycemia is defined as a BG level >10 mmol/L measured 60 – 120 minutes

after a meal. The Danish Health Authority advise patients with T1DM to measure one value of BG in the time interval of 60 – 120 minutes after a meal as part of the daily self-monitoring of blood glucose (SMBG) routine, but only if the preprandial BG level is normal and HbA1c is increased. [16] Proper detection of postprandial hyperglycemia is essential, but as the recommended interval for measuring postprandial BG is wide, and only one measurement is to be performed within the interval [16], the risk of missing a case of postprandial hyperglycemia is significant [17].

Continuous glucose monitoring (CGM) may improve detection rates, but CGM has not gained as large a foothold on the market of diabetes therapy as the more frequently SMBG routine, which is cheaper and requires less care. Thus, the timing of the one SMBG measurement is the key to how well post-prandial hyperglycemia can be detected. To our knowledge, only few studies have investigated how well suited the recommended interval is for detecting postprandial hyperglycemia in T1DM patients [18-21]. Across meals, only 67 – 71% of patients have been found to have blood glucose peak times within the recommended interval [19,20].

Therefore the aim of this study was to investigate whether, in cases of actual postprandial hyperglycemia, an alternative measurement interval would contain a higher percentage of hyperglycemic glucose levels than would the recommended interval, as the higher percentage would increase the likelihood of detecting a hyperglycemic value when doing only one glucose sample.

Methods

Study Population and Design

In this retrospective study, CGM datasets from a clinical development phase of the SCGM 1 system (Roche Diagnostics, Mannheim, Germany) were included. The data were collected at four medical centers (Medical Department M, Aarhus University Hospital, Aarhus, Denmark; Profil Institute for Metabolic Research, Neuss, Germany; German Diabetes Research Institute at the Heinrich–Heine University of Dusseldorf, Germany; and Department of Pharmaceutical Technology and Biopharmacy, University Center of Pharmacy, University of Groningen, the Netherlands). The system recorded glucose concentrations every minute for up to 5 days. For each experiment, the first 18 hours of data were discarded.

The datasets contained information on time and amount of carbohydrate intake and insulin injections. During the experiment the patients were not given access to the CGM data [9].

The study was performed by accessing the datasets of 159 patients. Only patients with T1DM were included. Patients were excluded if they experienced sudden illness during the experiment, if no events (carbohydrate intake or insulin injections) were registered or if continuous subcutaneous insulin infusion was used. After application of these restrictions, 51 patients were included for further analysis. After calibration and meal selection, which is described below, a total of 23 patients were included in the study. Patients were excluded if they had no valid meals using the criteriae below.

Calibration

In order to calibrate the CGM values, SMBG glucose measurements from a finger-capillary were obtained by a nurse up to 20 times a day using a blood glucose meter. Due to CGM inaccuracies, a linear regression approach was used to calibrate the CGM with SMBG measurements [22,23]. For each patient, a maximum of four SMBG measurements per 24-hours were used for calibration. The four measurements were found by using the first SMBG measurement after the discarded 18 hours, and then the next measurement closest to 6 hours after, and so forth.

Due to a physiological delay from plasma to interstitial glucose levels, the CGM data was delayed 10 minutes before pairing them with the SMBG values [23]. In case of two reference SMBG measurements at the same minute, they were averaged before use. Before a CGM-SMBG pair was made, the CGM value was screened in order to reject measurements of poor quality. Firstly and due to the physiological limits of blood glucose, only CGM values in the range of 40 mg/dL – 400 mg/dL were considered valid [24]. CGM-SMBG pairs where the CGM value was outside this interval were discarded. Secondly and due to the limited physiological change rate of BG, a CGM value was considered valid only if it deviated less than 4 mg/dL/min compared with the previous and the following value [24,25]. In case of a CGM value difference above 4 mg/dL/min, the CGM-SMBG pair was excluded and replaced by the next available pair. Calibration was visualized in calibration graphs with a linear regression equation and goodness of fit (R^2). Patients were excluded if they did not have enough SMBG values for calibration, in case of signal drops >6 hours or if the patient had an R^2 <0.70.

Selection of Meals

In order to investigate BG levels for postprandial hyperglycemia, meals were selected based on the time of carbohydrate intake. Breakfast, lunch and dinner were defined as all meals in the time range from 5.00 am – 11.00 am, 11.01 am – 4.00 pm and 4.01 pm – 10.00 pm, respectively. The meal with the highest amount of carbohydrate ingestion within each interval was selected as the actual main meal of the type in question. Only clear, single meals were included, as meals were excluded if another meal was ingested in the time interval of 150 minutes before to 180 minutes after the meal. In addition, meals were excluded in case of signal drop ≥5 minutes within this time interval.

After selection of meals, the CGM values for the first 180 minutes after a meal were screened for outliers. A value was considered an outlier if the value increased or decreased more than 4 mg/dL/min. In such cases, the CGM value was replaced by the mean of the previous and the following valid CGM value. Then, moving averages of 5 minutes were calculated for each minute. Lastly, the CGM values were calculated to SMBG values with the calibration equation.

Statistics

Microsoft Excel (version 14.0.7190.5000, 2010, Microsoft, Seattle, Washington, USA) was used to process and analyze data. Each meal was controlled for postprandial hyperglycemic episodes, which was defined as a BG level ≥10 mmol/L for a consecutive period of minimum 15 minutes. If a postprandial hyperglycemic episode was found at any time within 180 minutes after ingestion of a meal, the meal was classified as a case of "true postprandial hyperglycemia" and thus used for further analysis. The recommended measure interval of 60 – 120 minutes and the alternative interval of 120 – 180 minutes were examined after each meal, and the time with BG levels were above 10 mmol/L was calculated. Afterwards, the time was converted to percentage for each meal at each interval and a total average value for all meals, and breakfast, lunch, and dinner, respectively, was calculated for each of the two intervals.

SPSS (version 24, IBM, SPSS®, USA) was used for statistical analyses. A Shapiro-Wilk test was used to test all data for normality. The data did not follow a normal distribution. Therefore, a Wilcoxon test was used to test whether the alternative interval had a higher percentage of time of postprandial hyperglycemia than the recommended interval. Data were considered statistically significant with a p-value <0.05.

Results

Demographic Characteristics

23 patients who participated in the clinical development phase of the SCGM 1 system were included in this retrospective study. After analyzing all meals, postprandial hyperglycemic episodes were found after 63 meals, and therefore they were included in the study. The 63 meals were distributed as follows; breakfast: 25, lunch: 19, and dinner: 19. Two of the patients did not have postprandial hyperglycemia and therefore, they were excluded from further analysis. The remaining 21 patients included 13 men and 8 women. The median age of the patients was 33 years (range 18-53). The patients had a duration of DM of 14.86 ± 12.19 years (mean ± SD), and an HbA1c value of 8.25 ± 1.71% (mean ± SD).

Postprandial Hyperglycemia after Breakfast, Lunch and Dinner

The time, a patient experienced postprandial hyperglycemia, was recorded in the two different intervals after a meal, and converted to percentage for each interval for breakfast, lunch and dinner, respectively. The intervals are referred to as interval 1 (the recommended interval; 60 – 120 minutes after the meal) and interval 2 (the alternative interval; 120 – 180

minutes after the meal). The median values of the percentage of time, the patients spent in postprandial hyperglycemia, suggested a difference between the intervals (*figure 1*). For all breakfast measurements, the percentage of time associated with postprandial hyperglycemia was higher at interval 1 compared to interval 2, with median values of 83.34% and 76.67%, respectively. Although, the values could indicate a difference between the intervals (p = 0.225), this was not statistically significant.

The opposite was observed for the lunch measurements with median values of 65.00% and 88.34% for interval 1 and 2, respectively. The difference between the intervals (p = 0.925) were, however, not statistically significant.

The median values for the dinner measurements were 85.00%, and 60.00% for interval 1 and 2, respectively. This indicated a difference between the intervals (p = 0.776), but the intervals were not statistical significantly different from each other.

Postprandial Hyperglycemia after All Meals

The median values of the percentage of the interval, the patients spent in postprandial hyperglycemia for all meals together, indicated a difference between the two intervals (figure 2). For all meals, the percentage of time in postprandial hyperglycemia was higher at interval 1 with a median value of 80.00% compared with a value of 76.67% for interval 2. However, the statistic test revealed no significant difference between the intervals (p = 0.376).
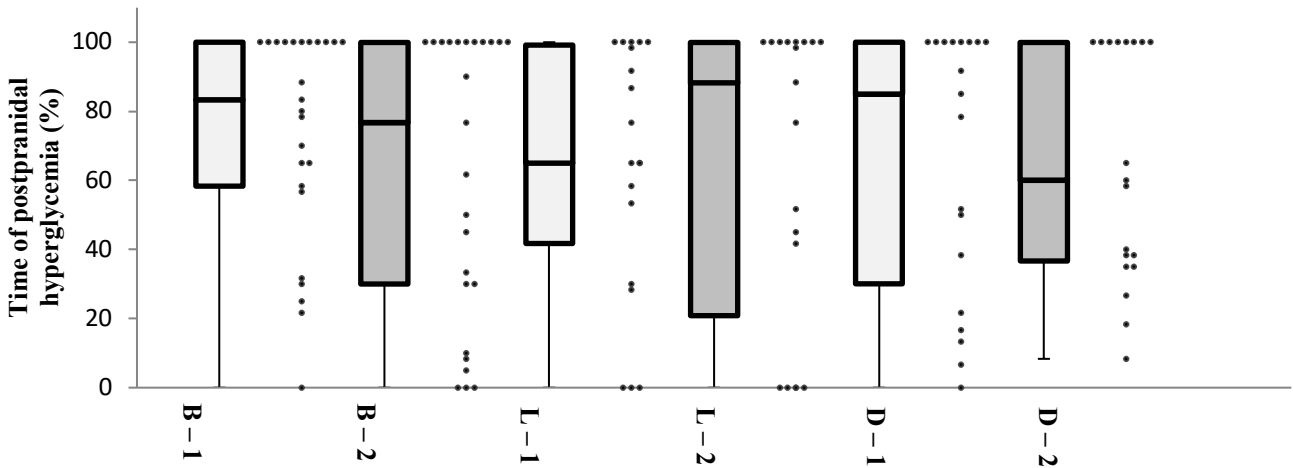

Figure 1- The amount of time spent in postprandial hyperglycemia in percentage (%) divided into breakfast (B), lunch (L), and dinner (D) for the two different intervals (median ± IQR). The data are named from the main meal (B: Breakfast, L: Lunch, D: Dinner) and the number of the interval (1 or 2). The colors represent the two different intervals, where BG levels were measured after a meal; light grey: interval 1 (60 – 120 mins.), and dark grey: interval 2 (120 – 180 mins.). The central rectangle spans the first quartile to the third quartile (IQR). The segment inside the rectangle shows the median and whiskers below the boxes show the minimum time associated with postprandial hyperglycemia. The dots represent the individual measurements for each patient stated for each meal and interval. %: percentage, IQR: interquartile range, BG: blood glucose.
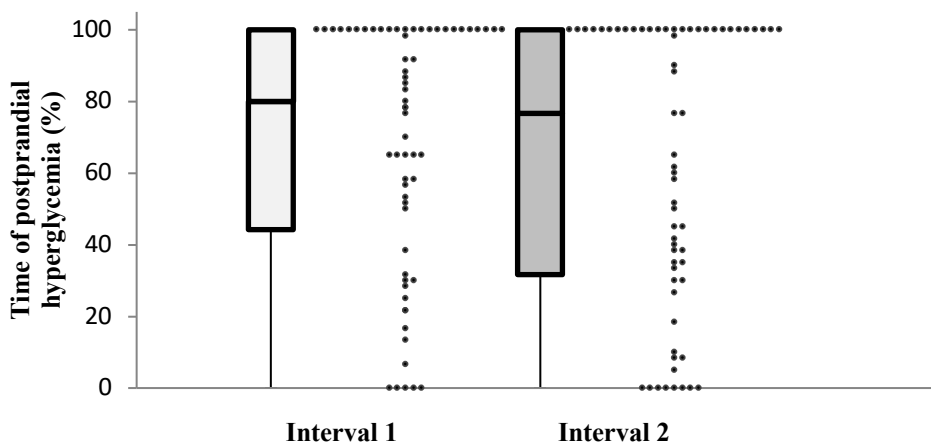
Figure 2-The amount of time spent in postprandial hyperglycemia in percentage (%) after all meals divided into the two intervals (median ± IQR). The colors represent the two different intervals, where BG levels were measured after a meal; light grey: interval 1 (60 – 120 mins.), and dark grey: interval 2 (120 – 180 mins.). The central rectangle spans the first quartile to the third quartile (IQR). The segment inside the rectangle shows the median and whiskers below the boxes show the minimum time associated with

postprandial hyperglycemia. The dots represent the data points for all meals at each interval. %: percentage, IQR: interquartile range, BG: blood glucose.

Discussion

This retrospective study investigated whether an alternative measurement time interval of 120 − 180 minutes after a meal would increase the likelihood of detecting postprandial hyperglycemia (>10 mmol/L) with postprandial SMBG, compared with the recommended interval of 60 − 120 minutes after a meal. This was evaluated based on the percentage of time the patients spent in hyperglycemia after a meal in the two intervals.

Findings from this study did not demonstrate any statistical significant differences between the recommended interval (interval 1) and the alternative interval (interval 2), which suggested that the alternative interval was not better for detecting postprandial hyperglycemia than the recommended interval and vice versa. Based on the median values, an indication of a higher percentage of time spent in postprandial hyperglycemia was however observed at interval 1 after breakfast ($p = 0.225$), dinner ($p = 0.776$), and when all meals were compared ($p = 0.376$). The opposite was observed after lunch, where an indication of a higher percentage of time spent in postprandial hyperglycemia was observed at interval 2 compared to interval 1 ($p = 0.925$).

Postprandial Hyperglycemia after All Main Meals

As postprandial hyperglycemia affects HbA1c and thereby also the risk of complications [6,9,13-15,26], it is important to focus on detection of postprandial hyperglycemia to ultimately minimize the occurrence.

Previously studies investigated mainly the timing of SMBG to detect postprandial hyperglycemia [18-21]. These findings are not easily compared with the findings of this study, as only peak times and peak values were reported. However, one of the studies also reported mean increase and decrease rate, which makes it possible to calculate the duration of postprandial hyperglycemia [20].

The peak times were reported to be within 57 - 100 minutes after breakfast, lunch and dinner [18,20,21], but the duration of postprandial hyperglycemia might be lengthy since a mean duration of a postprandial hyperglycemic episode was found to be 2.3 ± 1.1 hours (mean ± SD) for T1DM patients [27]. Daenen et al. found durations of postprandial hyperglycemia to last from 37 − 126 minutes, from 81 to 100 minutes and from 70 to 106 minutes after breakfast, lunch and dinner, respectively [20]. These were converted to percentages of time spent in postprandial hyperglycemia for the two intervals, which resulted in the highest percentage at interval 1 after breakfast, lunch and dinner. The percentage of time spend in postprandial hyperglycemia was 0% for interval 2 for all meals except from breakfast where it was 9.48%. The finding of interval 1 as the better interval for detection of postprandial hyperglycemia than interval 2 is consistent with the results of the present study for breakfast, dinner and when all meals were compared. Although for lunch this study found an indication of

interval 2 as a better time to measure BG to detect postprandial hyperglycemia.

Similar findings were observed when all meals were investigated together. Studies reported that 67 - 71% of the patients' peak times were within the recommended interval 1 [19,20], whereas 0% were within interval 2 [20]. These findings are consistent with the indications found in this study, as the longer duration is expected to be found around the peak time.

Generally, the findings of this study indicate that the detection rate is probably better in the recommended interval (interval 1). Still, the risk of missing postprandial hyperglycemia remains a challenge as the patients, when all meals are investigated together, did not have postprandial hyperglycemia in 20.00% (12 minutes) of this interval. In addition, it may be difficult to recommend one interval for all T1DM patients, as studies have reported large inter-individual differences [18-20] and as factors such as gender, HbA1c and DM duration might affect postprandial hyperglycemia.

Limitations

This study has limitations that need to be considered before drawing any firm conclusions. The study was a retrospective study and this may introduce some bias in the outcome. Generally, prospective cohorts that allow inclusion of subjects for the actual purpose, yield more reliable results. The relatively small sample size and the low number of meals might influence the power of this study. Preprandial BG levels, HbA1c, gender and duration of DM was not considered and could have affected the outcomes. Furthermore, data were obtained in Germany, Denmark and the Netherlands, and there may be large variations in dietary habits across the countries. Lastly, this study was performed in a hospitalization setting, and therefore the results may not truly reflect the everyday lives of T1DM patients.

Conclusion

This study found no statistical significant difference between the recommended interval and the alternative interval among postprandial hyperglycemia after all main meals in T1DM patients. An indication of a larger percentage of time spent in postprandial hyperglycemia was however observed when BG levels were measured at the recommended interval of 60 − 120 minutes after breakfast, and dinner, and when all meals were compared. This indicates that the recommended interval is a better time to measure BG levels for detecting postprandial hyperglycemia than the alternative interval. Further research in a larger group is however needed to validate the indications found in this study.

References

[1] IDF Diabetes Atlas 2017, 8th edition. Brussels, Belgium.: International Diabetes Federation 2017.

[2] Daneman D. Type 1 diabetes. Lancet 2006; 367: 847-858.

[3] Global report on diabetes 2016. Switzerland.: World Health Organization 2016.

[4] Vieira A, Druelle N, Avolio F, Napolitano T, Navarro-Sanz S, Silvano S, et al. Beta-Cell Replacement Strategies: The Increasing Need for a "beta-Cell Dogma". Front Genet 2017; 8: 75.

[5] Brod M, Nikolajsen A, Weatherall J, Pfeiffer KM. Understanding Post-Prandial Hyperglycemia in Patients with Type 1 and Type 2 Diabetes: A Web-based Survey in Germany, the UK, and USA. Diabetes Ther 2016; 7(2): 335-348.

[6] International Diabetes Federation Guideline Development Group. Guideline for management of postmeal glucose in diabetes. Diabetes Res Clin Prac. 2014; 103(2): 256-268.

[7] Ketema EB, Kibret KT. Correlation of fasting and postprandial plasma glucose with HbA1c in assessing glycemic control; systematic review and meta-analysis. Arch Public Health 2015; 73: 43

[8] Pfeiffer KM, Sandberg A, Nikolajsen A, Brod M. Postprandial glucose and healthcare resource use: a cross-sectional survey of adults with diabetes treated with basal-bolus insulin. J Med Econ 2018; 21(1): 66-73.

[9] Nielsen JK, Gravholt CH, Djurhuus CB, Brandt D, Becker J, Heinemann L, et al. Continuous subcutaneous glucose monitoring shows a close correlation between mean glucose and time spent in hyperglycemia and hemoglobin A1c. J Diabetes Sci Technol 2007; 1(6): 857-863.

[10] Writing Group for the DCCT/EDIC Research Group, Orchard TJ, Nathan DM, Zinman B, Cleary P, Brillon D, et al. Association between 7 years of intensive treatment of type 1 diabetes and long-term mortality. JAMA 2015; 313(1): 45-53.

[11] McCarthy M. American Diabetes Association issues new guidelines for type 1 diabetes. BMJ 2014; 348: g4119.

[12] Martin CL, Albers JW, Pop-Busui R, DCCT/EDIC Research Group. Neuropathy and related findings in the diabetes control and complications trial/epidemiology of diabetes interventions and complications study. Diabetes Care 2014; 37(1): 31-38.

[13] Madsbad S. Impact of postprandial glucose control on diabetes-related complications: How is the evidence evolving? J Diabetes Complications 2016; 30(2): 374-385.

[14] Gingras V, Taleb N, Roy-Fleming A, Legault L, Rabasa-Lhoret R. The challenges of achieving postprandial glucose control using closed-loop systems in patients with type 1 diabetes. Diabetes Obes Metab 2018; 20(2): 245-256.

[15] Ceriello A. The glucose triad and its role in comprehensive glycaemic control: current status, future management. Int J Clin Pract 2010; 64(12): 1705-1711.

[16] Kliniske retningslinier for behandling af voksne med type 1 diabetes 2010. Denmark.: Dansk Endokrinologisk Selskab 2010.

[17] Bergenstal RM, Ahmann AJ, Bailey T, Beck RW, Bissen J, Buckingham B, et al. Recommendations for standardizing glucose reporting and analysis to optimize clinical decision making in diabetes: the Ambulatory Glucose Profile (AGP). Diabetes Technol Ther 2013; 15(3): 198-211.

[18] Taki K, Nishimura R, Morimoto A, Tsujino D, Miyashita Y, Tajima N. Analysis of 24-hour glycemic excursions in patients with type 1 diabetes by using continuous glucose monitoring. Diabetes Technol Ther 2010; 12(7): 523-528.

[19] Johansen MD, Gjerlov I, Christiansen JS, Hejlesen OK. Interindividual and intraindividual variations in postprandial glycemia peak time complicate precise recommendations for self-monitoring of glucose in persons with type 1 diabetes mellitus. J Diabetes Sci Technol 2012; 6(2): 356-361.

[20] Daenen S, Sola-Gazagnes A, M'Bemba J, Dorange-Breillard C, Defer F, Elgrably F, et al. Peak-time determination of post-meal glucose excursions in insulin-treated diabetic patients. Diabetes Metab 2010; 36(2): 165-169.

[21] Manuel-y-Keenoy B, Vertommen J, Abrams P, Van Gaal L, De Leeuw I, Messeri D, et al. Postprandial glucose monitoring in type 1 diabetes mellitus: use of a continuous subcutaneous monitoring device. Diabetes Metab Res Rev 2004; 20(Suppl 2): 24-31.

[22] Danne T, Nimri R, Battelino T, Bergenstal RM, Close KL, DeVries JH, et al. International Consensus on Use of Continuous Glucose Monitoring. Diabetes Care 2017; 40(12): 1631-1640.

[23] Cengiz E, Tamborlane WV. A tale of two compartments: interstitial versus blood glucose monitoring. Diabetes Technol Ther 2009; 11(Suppl 1): 11-16.

[24] Mahmoudi Z, Dencker Johansen M, Christiansen JS, Hejlesen OK. A multistep algorithm for processing and calibration of microdialysis continuous glucose monitoring data. Diabetes Technol Ther 2013; 15(10): 825-835.

[25] Facchinetti A, Sparacino G, Cobelli C. An online self-tunable method to denoise CGM sensor data. IEEE Trans Biomed Eng 2010; 57(3): 634-641.

[26] Jiang J, Zhao L, Lin L, Gui M, Aleteng Q, Wu B, et al. Postprandial Blood Glucose Outweighs Fasting Blood Glucose and HbA1c in screening Coronary Heart Disease. Sci Rep 2017; 7(1): 14212-017-14152-y.

[27] Bode BW, Schwartz S, Stubbs HA, Block JE. Glycemic characteristics in continuously monitored patients with type 1 and type 2 diabetes: normative values. Diabetes Care 2005; 28(10): 2361-2366.

Address for correspondence

Mia Birkholm Lausten

MSc in Medicine with Industrial Specialization

Email: Mia.lausten@hotmail.com

An International Minimal Patient Care Report Exemplified in FHIR to Facilitate Standardisation and Interoperability of Emergency Medical Services Data

Rasmus Guldhammer Blendal, Louise Pape-Haugaard

Department of Health Science and Technology, Aalborg University, Aalborg, Denmark

Abstract

Lack of semantic interoperability is a common problem within healthcare IT. In prehospital sector it results in emergency medical service data being stuck in the emergency department, despite being important to clinicians outside the emergency department in treatment of the patient. Steps towards semantic interoperability within the prehospital area, on an international level, was taken by first creating an international common data set by use of two different national data sets and scientific literature. The usability of the created data set was showcased by example of a clinical case from emergency medicine. With common data elements established the next step was profiling the data set in FHIR, to further facilitate interoperability, as a common exchange standard is paramount for systems to interoperate.

Keywords:

Interoperability, Emergency Medical Services, Emergency Department, FHIR, Standardization.

Introduction

The idea of collecting patient data before patients reach the hospital, has been around for half a century in the US [1]. In 1966 a study investigated how lack of prehospital data could have an influence on hospital acquired mortality [1]. In 1995 consensus on a standardised comprehensive list of data points to be collected from trauma patients was reached. This standardised list of data points is the basis for the American National Emergency Medical Services Information Systems (NEMSIS), which is used across the US to store and standardise data collected by Emergency Medical Service (EMS) agencies. NEMSIS enables the possibility to check performance of clinical interventions in the prehospital field and conduct cohort studies [2]. However, in European eHealth consisting of different nations, different legislations, different vendors the requirements for prehospital data points have a high degree of variance and have created various data sets. A possible solution to proprietary data sets and lack of interoperability, is to standardise the data collected by EMS internationally along with the way the EMS data is exchanged internationally. A such solution was made by McClay et. Al. [3] called DEEDS, a purely US centric data set with 725 data elements, included in the Health Level 7 (HL7) v3 specification. A study by Brammen et. Al. [4] compared the data elements in DEEDS with the German equivalent (GEDMR). The authors found significant differences between the two data sets in certain domains and their results support the development of an international standard for ED data elements [4]. An issue with DEEDS and GEDMR is the sheer size of the data sets which can limit adaptation, due to the work required to map the data elements to existing ones in the individual EHR systems. Additionally, DEEDS is part of the HL7 v3 specification, a very robust but ill-adopted standard due the extensive implementation work required, why it did not see much adoption worldwide compared to its predecessor. A solution to the issues with DEEDS and GEDMR is to create a minimal data set and use a specification that allows for simpler implementation. An emerging standard for sharing healthcare data across systems is the Fast Healthcare Interoperability Resources (FHIR) from HL7 [5]. FHIR enables standardisation of clinical data exchange, by allowing context to be kept, and terminology bound to the data [5, 6].

The objective in this study is to investigate how sharing of clinical meaningful minimal EMS data can be facilitated by the use of FHIR.

Materials and Methods

To investigate how sharing of clinical meaningful minimal EMS data can be facilitated by the use of FHIR we need to agree on a common international data set. Hence, the first step is to create such a data set. Further, we need to investigate how this data set can be profiled using FHIR. To profile in FHIR we have identified a clinical case.

The International Data set

The international data set consists of the intersection data points from two national standardised lists of data points. Variations between nations in collected data was expected, why the minimal data set will only cover the most clinically essential data for the emergency department (ED) to properly treat the patient. Examples of intersecting categories in EMS data are: Injury information [7], medical history [8], allergies [7], procedures performed [7], and patient physiology measures [9]. Additionally, information in relation to the

patient's symptoms/ immediate diagnosis and triage levels to determine the amount of personnel and resources required for the given patient [10]. The main requirements for a data element to be included in the minimal data set are: a) clinically essential, as previously defined, for the ED to understand the prehospital patients' situation, b) essential for the ED facilitate care of the patient. In this study, we deselect logistical and administrative data points.

Elements present in both EMS data sets but not fulfilling the clinical relevance requirements will not be used. When there is consensus on the category, but too little overlap between the data elements used to describe the category, literature will be used instead of the two national EMS data sets to find clinically relevant data elements.

Profiling the clinical case in FHIR

An international minimal data set for the prehospital sector, from here on referred to as IMPCR (International Minimal Patient Care Report), can only solve part of the problem described previously. The technical part of interoperability has to be solved using a standard exchange protocol fit for clinical data exchange. FHIR was chosen as the communication standard due its rapid adoption by vendors and an active community, making it a very promising standard that can see widespread use. Further, international projects have begun working towards international interoperability between EHR systems using FHIR, such as the Trillium project [11]. The clinical case chosen to be profiled in FHIR is presented in table 1.

Tabel 1- The clinical case has been adapted from [12]

Clinical case
The local EM services have been called to the local riding club to collect Kathrine Smith, a 24-year old woman who has fallen off her horse. Katherine is conscious but has pain in her pelvic area, she rates it an 8 out of 10. The EMS personnel notes that she has no known allergies and is an otherwise healthy young woman. Due to the height of the fall, horse was in mid jump, she is put in a hard collar and taped to the hard spine board to ensure spinal stability. Once in the ambulance her vitals are measured Temp: 36.2C, HR: 90/min, RR: 22/min, BP: 110/ 65 mmHg, SpO2: 98%OA and GCS 15. She is given a 50 mcg bolus of Fentanyl for the pain along with an additional 75 mcg via IV during the ambulance ride. The iliac crest feels tender when palpated. The EMS personnel suspects a pelvic fracture, performs a triage and transport her to the ED.

Results

The International Data set

The categories and data elements of IMPCR can be seen in Table 2. The four categories marked with an asterisk contains data elements from the literature. The injury category takes its elements from trauma registries in Europe and USA such

as [13, 14], to ensure data collected can be transferred to appropriate registries. Triage uses the Emergency Severity Index v4 as both national EMS data sets used a proprietary triage system, instead of developing a third triage system based upon the two, an internationally validated triage system was chosen [15, 16, 17]. Patient Assessment uses the ABCDE system which has widespread use and is recommended by the European Resuscitation Council [18, 19]. Cardiac arrest was another category where the two ePCRs differed, a report form with international backing and consensus from multiple cardiac societies was therefore chosen [20]. The remaining categories and data elements were found in the two EMS data sets.

Table 2 - Categories and data elements of the International Minimal Patient Care Report. The asterisk denotes the elements were found in the literature

International Minimal Patient Care Report (IMCPR)	
Patient	Name Gender DOB Identifier
Additional Patient Information	Allergy Severity Recent surgery/medical history Time of surgery/medical history
Medication	Active Ingredient Dosage Unit Dispensation
Cardiac Arrest*	Cardiac arrest determined by: Cause of arrest, Treatment before EMS arrival, Bystander CPR Defibrillation before EMS Resuscitation attempted by EMS Locations of arrest Witnessed Initial rhythm Chest impressions Ventilation Time of: collapse, call receipt, vehicle stopped, first rhythm analysis, death or cessation of CPR Spontaneous circulation on arrival
Injury*	Type Mechanisms Traffic factors
Diagnose/symptoms	Central Nervous System Respiratory System Musculoskeletal + Connective Tissue Circulatory System Abdomen Psychiatric/Mental/Behavioural Other Primary/action diagnosis
Patient assessment	Airways Breathing

	Circulation
	Disability
	Exposure
Procedures	Time of procedure
	Procedure performed
	Performed by
Vital signs	ECG (measurement, Type, interpretation)
	Systolic BP
	Diastolic BP
	Heart Rate
	SpO_2
	Respiration Rate
	$ETCO_2$
	Blood Glucose Level
	GCS total + Sub scores
	Pain Scale
	Body Temperature
Triage	Requires immediate lifesaving intervention?
	High risk situation?
	Confused/lethargic/disoriented?
	Severe pain/distress?
	How many different resources are needed?
	Danger zone vitals?
	ESI Score

Profiling the clinical case in FHIR

Patient: Allergy information was profiled in the same way as medical history, with one code/value indicating "No known allergies", since the patient had no allergies. The profiled Patient was very similar to the base resource, as only animal and multipleBirths had their cardinality set to 0..0 as these elements are never relevant, while gender and DOB had their cardinality set to 1..1 to emphasise their importance.

Medication: To profile Medication, MedicationAdministration was chosen, as this resource specifies what, how, and by whom medication was administered. Few FHIR elements had their cardinality set to 0..0. The elements with their cardinality set to 0..0 were: 'notGiven', 'reasonNotGiven', 'prescription', 'device', and 'eventHistory'. These were removed because they are not needed for the prehospital use-case. The prescription covers "the original request/instruction or authority to perform the administration", which is not relevant for EM services, as EMS personnel are usually not allowed to administer medication not within their authority, why performer will be sufficient to track who administered the medication.

Injury: Injury was profiled using the Condition resource, as the injury itself has a direct impact on the patient's health. Specifying the type of injury and the mechanisms behind it, gives insight into the patient's condition and the required resources needed to treat the patient in the ED. Only two elements in Condition are required to capture the injury information from the clinical case, why rest of the elements are left as is. Both 'category' used to specify the type of injury and 'code' used to specify the mechanism behind the injury

are sliced, instead of using valuesets. Slicing was used to illustrate what the valueset should contain, each slice holds a SNOMEDCT value, that has been chosen without further thought, as they are strictly example codes.

Diagnosis/Symptoms: Diagnose/Symptoms was modelled in FHIR using Condition. The main elements used in the resource are 'category', 'severity' and 'code'. category should ideally contain a valueset which holds the data elements of Diagnose/Symptoms e.g. Central Nervous System and Respiratory System. code should ideally be a valueset for the chosen category e.g. for respiratory system it could be: Pneumonia, Pneumothorax. The primary diagnosis can be noted in severity.

Vital signs: All measurements but ETCO2 and blood glucose were existing profiles found at https://registry.fhir.org/. ETCO2 and blood glucose were profiled by copying a heart rate profile and changing the 'names' and 'code'. Glasgow Coma Scale was profiled by having the total score as the value of the observation and the three sub scores as components of the total value.

Triage: The ESI Triage model can be seen as a questionnaire with a series of yes/no question and a resulting score. The questionnaire was profiled using the questionnaire resource by slicing 'item' and creating one less number of slices than number of questions. Because question A: Requires immediate lifesaving intervention? is contained in the outermost item and therefore not part of the slices. Next, each question had the answer set linked to an existing valueset if possible, valuesets for Yes/No did exist. In the case the answer set did not exist as a valueset the answer element was sliced to contain the custom answers (e.g: None/One/Many).

Procedure: Procedure was modelled using Procedure, as it seemed the obvious choice. The procedure, which should contain a valueset of possible procedures, was modelled to 'code' in the resource, with four example procedures as slices. The profile is able to store the expertise level of the performer such as: basic life support, advanced life support or physician. Information that might be recorded automatically system side.

The clinical case is modelled in FHIR and presented in Table 3.

Tabel 3- Illustrates how the clinical case is modelled in FHIR

Clinical Case	IMPCR Category	FHIR Resource(s)
...Kathrine Smith, a 24-year old woman The EMS personnel notes that she has no known allergies and is an otherwise healthy young woman	Patient + Additional patient information	Patient, AllergyIntolerence, Observation

..fallen off her horse...horse was in mid jump..	Injury	Observation
..is conscious but has pain in her pelvic area, she rates it an 8 out of 10.	Vital signs + Diagnose/ symptoms	Observation.Vitalsigns
...put in a hard collar and taped to the hard spine board to ensure spinal stability. . .	Procedure	Procedure
Once in the ambulance her vitals are measured Temp: 36.2C, HR: 90/min, RR: 22/min, BP: 110/ 65 mmHg, SpO2: 98%OA and GCS 15	Vital signs	Observation.Vitalsigns
She is given a 50 mcg bolus of Fentanyl for the pain along with an additional 75 mcg via IV during the ambulance ride	Medication	MedicationAdministration
The iliac crest feels tender when palpated. The EMS personnel suspects a pelvic fracture...	Diagnose/symptoms	Condition
..Performs a triage..	Triage	Questionnaire

Discussion

The International Data set

The national EMS data set in USA and a small European country have been chosen due to the difference between the two countries healthcare systems, insurance/publicly funded vs entirely publicly funded respectively. The choice of using national EMS data sets from different healthcare systems will increase the likelihood of identifying the relevant data points to include in the new minimal data set, as the EMS data sets are expected to vary more than that of two homogeneous healthcare systems. As the aim of the data set was to be minimal, coarser granularity were used when possible, to reduce the amount of data points. The benefit of an intermediate level of granularity is faster data retrieval for the clinician, as they try to get an overview of the patient [21]. An example of this is the disease/diagnose category, where the intended value sets, based on the two national data sets, are the most common diseases/symptoms encountered in the prehospital sector split into organ systems. An alternative design would

be to register the overall category, then the exact anatomical location and then the symptom. Instead the coarser version was chosen, where the diagnosis is registered to reduce the data point having to be collected. Location of the disease/symptom can in some instances also be inferred based upon the organ system and diagnose/symptom chosen.

Two national data sets and literature was used to create IMPCR, a data set containing clinical data elements able to describe the prehospital patient and their situation. IMPCR is mostly based on the two data sets, as it gave a solid departure point for identifying the required clinical categories. If consensus on data elements can be found between two data sets independently developed, the element must have clinical relevance. An additional data set could have been added, it might have had the benefit of settling disagreements, further confirming consensus, or in the worst case added further disagreement. In case of more disagreement more literature would have been used if possible. As this clinical data sets builds upon existing data sets and aggregates clinical report forms from different clinical specialities, it does not follow the usual method of developing a clinical data set, it builds upon existing work instead of starting from scratch. This is an advantage because the data sets are already being used, thus have a proven track record of being able to provide the adequate information to clinicians in the ED. The addition of validated report forms from clinical societies, results in further standardisation of how clinical data is reported and can ideally support research in addition to the clinical use case. Other data sets have been created or revised using systematic literature reviews, expert opinion and consensus found via techniques such as 'Nominal Group Technique' [13, 22].

IMPCR's capabilities were tested using a realistic clinical case, see table 6.1, where the clinical case was split into its components and matched with the appropriate IMPCR category. IMPCR was able to capture the relevant data from the clinical case satisfactorily, however more cases, especially complicated ones, should be tried in the future. IMPCR is capable of capturing data meant for the clinic and by being a standardised data set, it enables comparison of emergency departments and emergency medical services both nationally and internationally. Using identical data sets allows for comparison because the data elements are directly comparable opposed to having to map from one data set to another or pick out a few characteristics [23] - which can bring benefit for research and quality control purposes. But having been creating with minimalism in mind, there is a risk it is not entirely suitable for research at a later time, this could be both due to granularity of the data or optionality. There is a risk of incomplete data when certain elements might not get collected if they are not applicable to the patient in question [24]. A potential issue that can be solved by removing optionality and requiring negative values when the category/element is not relevant. The current version of IMPCR additionally lacks terminology binding to properly facilitate semantic interoperability. Terminology binding is crucial as it keeps the semantic meaning of data instead of leaving it as an arbitrary value. Further it combines the multiple ways of describing a single medical concept into one, allowing systems not using the same data set, to understand what they

receive so long as they can look up the terminology. [25, 26, 27].

Flexibility of FHIR

FHIR being based on base resources that can satisfy a range of use-cases, can result in model variability and multiple right ways to model. An example of this was encountered when profiling medical history, there was a choice whether it should have been profiled as an Observation or a Condition, as both are able to hold the required data. Either choice can therefore work depending on the light in which the prior medical issues are viewed: as potential problems that will require additional medical attention, or as 'need to know' information for the emergency personnel. This way of having multiple ways of describing/modelling the same thing is possible because FHIR follows no standard model, opposed to more rigid standards such as HL7v3 which used the RIM [28]. The flexibility of FHIR means it alone is not sufficient to facilitate semantic interoperability, as there is a risk of independent project/vendors developing differing proprietary profiles for the same use-case, thus contributing to fragmentation and interoperability issues within the healthcare sector [29, 30].

As a response to this concern the SMART team [30], put forth a small set of profiles where they specify terminologies, element cardinality and type choice to ensure apps created with SMART, could semantically interoperate. Likewise, a larger scale project, Argonaut, now collaborating with SMART, has gathered traction [31, 32]. Argonaut is a collaboration between private vendors, such as Cerner and Epic, that are "Creating open industry Implementation Guides in high priority use cases of importance to patients, providers and the industry as a whole" [31], a project that is open to other organisations to foster interoperability [32]. Reusing profiles and sharing implementation guides in FHIR is best practice, to ensure widespread support of commonly used profiles, but this can result in first movers dictating the profiles. There is the risk of using sup-optimal profiles for the given resources, as first movers can create a tradition, and effectively take over the market. While interoperability is the goal of FHIR, an interesting discussion arises: On whose terms, should the goal be reached?

References

[1] D. Spaite, R. Benoit, D. Brown, R. Cales, D. Dawson, C. Glass, C. Kaufmann, D. Pollock, S. Ryan, and E. M. Yano, "Uniform prehospital data elements and definitions: a report from the uniform prehospital emergency medical services data conference," Annals of emergency medicine, vol. 25, no. 4, pp. 525–534, 1995.

[2] K. L. Tataris, M. P. Mercer, and P. Govindarajan, "Prehospital aspirin administration for acute coronary syndrome (ACS) in the USA: an EMS quality assessment using the NEMSIS 2011 database," Emergency Medicine Journal, vol. 32, no. 11, pp. 876–881, 2015.

[3] James C McClay, Peter J Park, Mark G Janczewski, Laura Heermann Langford, "Standard for improving emergency information interoperability: the HL7 data elements for emergency department systems", Journal of the American Medical Informatics Association, Volume 22, Issue 3, 1 May 2015, Pages 529–535

[4] Brammen D, Eggert P, Lucas B, Heermann-Langford L, McClay JC., "Comparing the German Emergency Department Medical Record with the US HL7 Data Elements for Emergency Department Systems",, Studies in health technology and informatics. 2018;247:216-20.

[5] "Fhir overview." http://www.hl7.org/implement/standards/fhir/overview.html . Accessed: 2018-03-21.

[6] M. Smits, E. Kramer, M. Harthoorn, and R. Cornet, "A comparison of two Detailed Clinical Model representations: FHIR and CDA," European Journal for Biomedical Informatics, vol. 11, no. 2, p. 11, 2015.

[7] J. P. Benner, J. Hilton, G. Carr, K. Robbins, R. C. Schutt, M. P. Borloz, K. Alibertis, B. Sojka, K. Hudson, D. Haugh, and W. Brady, "Information transfer from prehospital to ED healthcare providers," The American Journal of Emergency Medicine, vol. 26, pp. 233–235, Feb. 2008.

[8] J. T. Finnell and J. M. Overhage, "Emergency Medical Services: The Frontier in Health Information Exchange," AMIA Annual Symposium Proceedings, pp. 222–226, 2010.

[9] D. J. Laudermilch, M. A. Schiff, A. B. Nathens, and M. R. Rosengart, "Lack of Emergency Medical Services Documentation Is Associated with Poor Patient Outcomes: A Validation of Audit Filters for Prehospital Trauma Care," Journal of the American College of Surgeons, vol. 210, pp. 220–227, Feb. 2010.

[10] M. Christ, F. Grossmann, D. Winter, R. Bingisser, and E. Platz, "Modern Triage in the Emergency Department," Deutsches Aerzteblatt Online, Dec. 2010. [6] S. R. Bruijns, H. R. Guly, O. Bouamra, F. Lecky, and L. A. Wallis, "The value of the difference between ed and prehospital vital signs in predicting outcome in trauma," Emergency Medicine Journal, vol. 31, no. 7, pp. 579–582, 2014.

[11] "Trillium project." https://simplifier.net/TrilliumII. Accessed: 2018-04-12.

[12] C. Nickson, "Trauma! pelvic fractures i." https://lifeinthefastlane.com/trauma-tribulation-027/. Accessed: 2018-05-01

[13] K. G. Ringdal, T. J. Coats, R. Lefering, S. Di Bartolomeo, P. A. Steen, O. Roise, L. Handolin, H. M. Lossius, and U. T. e. p. (Tcd), "The Utstein template for uniform reporting of data following major trauma: a joint revision by SCANTEM, TARN, DGU-TR and RITG," Scandinavian Journal of Trauma, Resuscitation and Emergency Medicine, vol. 16, no. 1, p. 7, 2008.

[14] H. Champion, W. S. Copes, and W. J. Sacco, "The Major Trauma Outcome Study: establishing national norms for trauma care." The Journal of Trauma, vol. 30, pp. 1356–1365, Nov. 1990.

[15] N. Gilboy, P. Tanabe, and D. A. Travers, "The emergency severity index version 4: Changes to esi level 1 and pediatric fever criteria," Journal of Emergency Nursing, vol. 31, no. 4, pp. 357 – 362, 2005

[16] M. Christ, F. Grossmann, D. Winter, R. Bingisser, and E. Platz, "Modern Triage in the Emergency Department," Deutsches Aerzteblatt Online, Dec. 2010.

[17] N. A. Green, Y. Durani, D. Brecher, A. DePiero, J. Loiselle, and M. Attia, "Emergency Severity Index Version 4: A Valid and Reliable Tool in Pediatric Emergency Department Triage," Pediatric Emergency Care, vol. 28, pp. 753–757, Aug. 2012.

[18] T. Thim, Krarup, Grove, Rohde, and Lofgren, "Initial assessment and treatment with the Airway, Breathing, Circulation, Disability, Exposure (ABCDE) approach," International Journal of General Medicine, p. 117, Jan. 2012.

[19] J. Soar, J. P. Nolan, B. W. Böttiger, G. D. Perkins, C. Lott, P. Carli, T. Pellis, C. Sandroni, M. B. Skrifvars, G. B. Smith, K. Sunde, C. D. Deakin, R. W. Koster, K. G. Monsieurs, and N. I. Nikolaou, "European Resuscitation Council Guidelines for Resuscitation 2015," Resuscitation, vol. 95, pp. 100–147, Oct. 2015.

[20] I. Jacobs, V. Nadkarni, J. Bahr, R. A. Berg, J. E. Billi, L. Bossaert, P. Cassan, A. Coovadia, K. D'Este, J. Finn, H. Halperin, A. Handley, J. Herlitz, R. Hickey, A. Idris, W. Kloeck, G. L. Larkin, M. E. Mancini, P. Mason, G. Mears, K. Monsieurs, W. Montgomery, P. Morley, G. Nichol, J. Nolan, K. Okada, J. Perlman, M. Shuster, P. A. Steen, F. Sterz, J. Tibballs, S. Timerman, T. Truitt, and D. Zideman, "Cardiac arrest and cardiopulmonary resuscitation outcome reports: update and simplification of the Utstein templates for resuscitation registries.," Resuscitation, vol. 63, pp. 233–249, Dec. 2004.

[21] H. J. Tange, H. C. Schouten, A. D. M. Kester, and A. Hasman, "The Granularity of Medical Narratives and Its Effect on the Speed and Completeness of Information Retrieval," Journal of the American Medical Informatics Association, vol. 5, pp. 571– 582, Nov. 1998

[22] S. E. Lamb, E. C. Jørstad-Stein, K. Hauer, C. Becker, and on behalf of the Prevention of Falls Network Europe and Outcomes Consensus Group, "Development of a Common Outcome Data Set for Fall Injury Prevention Trials: The Prevention of Falls Network Europe Consensus: PROFANE COMMON OUTCOME DATA SET," Journal of the American Geriatrics Society, vol. 53, pp. 1618–1622, Sept. 2005.

[23] M. Fischer, J. Kamp, L. Garcia-Castrillo Riesgo, I. Robertson-Steel, J. Overton, A. Ziemann, and T. Krafft, "Comparing emergency medical service systems—A project of the European Emergency Data (EED) Project," Resuscitation, vol. 82, pp. 285–293, Mar. 2011.

[24] W. R. Hersh, M. G. Weiner, P. J. Embi, J. R. Logan, P. R. Payne, E. V. Bernstam, H. P. Lehmann, G. Hripcsak, T. H. Hartzog, J. J. Cimino, and J. H. Saltz, "Caveats for the Use of Operational Electronic Health Record Data in Comparative Effectiveness Research:," Medical Care, vol. 51, pp. S30–S37, Aug. 2013.

[25] A. Ryan, "Towards semantic interoperability in healthcare: ontology mapping from SNOMED-CT to HL7 version 3," in Proceedings of the second Australasian workshop on Advances in ontologies-Volume 72, pp. 69–74, Australian Computer Society, Inc., 2006.

[26] C. N. Mead, "Data interchange standards in healthcare it-computable semantic interoperability: Now possible but still difficult. do we really need a better mousetrap?", Journal of Healthcare Information Management, vol. 20, no. 1, p. 71, 2006.

[27] W. T. Goossen, "Intelligent semantic interoperability: integrating knowledge, terminology and information models to support stroke care," Consumer-Centered Computer-Supported Care for Healthy People, p. 6, 2006.

[28] D. Bender and K. Sartipi, "HL7 FHIR: An Agile and RESTful approach to healthcare information exchange," Proceedings of the 26th IEEE International Symposium on Computer-Based Medical Systems, pp. 326–331, June 2013.

[29] H. Ulrich, A.-K. Kock, P. Duhm-Harbeck, J. K. Habermann, and J. Ingenerf, "Metadata Repository for Improved Data Sharing and Reuse Based on HL7 FHIR.," in MIE, pp. 162–166, 2016.

[30] J. C. Mandel, D. A. Kreda, K. D. Mandl, I. S. Kohane, and R. B. Ramoni, "SMART on FHIR: a standards-based, interoperable apps platform for electronic health records," Journal of the American Medical Informatics Association, vol. 23, pp. 899–908, Sept. 2016.

[31] Argonaut, "Argonaut project, presentation feb. 2017." https://www.hl7.org/documentcenter/public/calendarofevents/himss/2017/TheArgonautProjectandHL7FHIR.pdf, February 2017. Accessed: 2018-05-31.

[32] Argonaut, "Argonaut project." http://argonautwiki.hl7.org/index.php?title=Main_Page. Accessed: 2018-05-31.

Address for correspondence

Rasmus Guldhammer Blendal - rasmusblendal@hotmail.com

A Method for Reporting of Variance in Informal Care Pathways

John J Chelsom[a,b], Conceição Granja[c]

[a]*Seven Informatics Ltd, Banbury, United Kingdom*
[b] *Health Information Science, University of Victoria, Victoria, Canada*
[c] *Norwegian Centre for E-health Research, Tromsø, Norway*

Abstract

Care pathways are commonly used as a means to ensure the implementation of proven best practice in clinical care. They can be defined in an Electronic Health Record system using formal representations of template pathways which are then instantiated for specific patients, to guide clinical users through a sequence of tasks and actions. The open source cityEHR system uses HL7 CDA documents for the formal pathway definitions, but in practice we have found that some care pathways, and clinical protocols, can be implemented effectively using less formal mechanisms. In addition to the shared patient record itself, we find that the most important of these mechanisms are notification on recording of patient events, support for care teams, annotation of the record, and the clinician-centred 'In-Tray'. One important feature lost in this approach is the ability to report on the variance of the care provided to the patient from the best practice encapsulated in a more formally defined pathway. To address this, we describe a method for transforming a formally defined care pathway to a finite state machine representation and querying the patient record to generate a retrospective report on the variance of states recorded in the patient record.

Keywords:

Care pathways, Clinical communication, Patient-centred EHR, HL7 CDA, Finite State Machine, Variance Report.

Introduction

A care pathway is a step-by-step guide for the care or treatment of a patient, in a defined clinical context, usually following an evidence-based recommendation of best clinical practice in that context. The implementations of care pathways in Electronic Health Record (EHR) systems generally adopt a formal approach to the definition of template pathways, which are then instantiated with data from an individual patient record to create specific instances that guide the care of that patient [1].

The open source cityEHR system uses HL7 CDA documents [2] to provide the formal definition of care pathways and records the progress of the pathway as a series of CDA documents stored in the patient record [3]. For some implemen-

tations of the cityEHR we have observed that alternative, and less formal, methods are more effective in managing the care pathways. We define an informal care pathway as being a set of mechanisms, configured in the EHR, which guide clinical users in following a sequence of actions, without a predetermined or prescribed specification of that sequence.

These mechanisms use the patient centred record as the hub for communication between the care team and the patient, in the ways described by Ueckert et al [4] and White & Danis [5]. The use of the EHR as the facilitator for communications has been identified as one of the key issues in the patient-centred medical home (PCMH) model in the United States [6].

As an example, we refer in this paper to an implementation of cityEHR which supports the activities of a Fracture Prevention Service (FPS) in the UK. This system, called Elfin, has been developed at the Nuffield Department of Orthopaedics, Rheumatology and Musculoskeletal Sciences at the University of Oxford. Elfin has been running in Oxford since March 2015, and has seen live operation at two other NHS sites in England. The system is built on the cityEHR platform, with an information model developed specifically to address the requirements of the Fracture Liaison Service.

The first implementation of Elfin was designed with a formal representation of the care pathway, using HL7 CDA as the underlying data representation. The advantage of using HL7 CDA is that the care pathways are stored in the record in the same form as all other clinical documents [3]. The Care Pathway engine in cityEHR then uses the formal definitions to drive the sequence of actions undertaken by clinical users, where the outcome of each action is itself recorded as an HL7 CDA document (e.g., a form, letter, order, prescription, etc.).

After the first implementation, we discovered that a more flexible, less formal, approach could be used; it is this version that was adopted for live use and its key features are described in the following section. One consequence of this alternative approach is that there is no inherent mechanism for reporting on variance (or compliance) with the formally defined pathway. Reporting of variance has been identified by Wakamiya and Yamauchi as one of the key features of a conventional care pathway implementation [7], and hence is an important issue to address.

Our objective, therefore, was to define and implement a method for reporting on the variance between actions recorded in the patient record and the set of actions defined formally in a care pathway, where that pathway is applied retrospectively, rather than being used to prescribe the sequence of actions as they are performed. This method is based on the transformation of the pathway into a finite state machine, as proposed by Eshuis [8], followed by pattern matching and temporal analysis as proposed by Awad et al [9].

In this paper we describe our approach to implementing care pathways using informal mechanisms of patient-centred clinical communication within the cityEHR, the feedback we have gathered after experience with an operational EHR and a method for retrospective reporting of variance from a formally specified pathway.

Methods

The Formally Defined Care Pathway

An example care pathway is shown in Figure 1, depicted as an Activity Diagram, following the conventions of the Unified Modelling Language (UML). This example is a simplified version of the Fracture Liaison Service pathway implemented in the Elfin system.
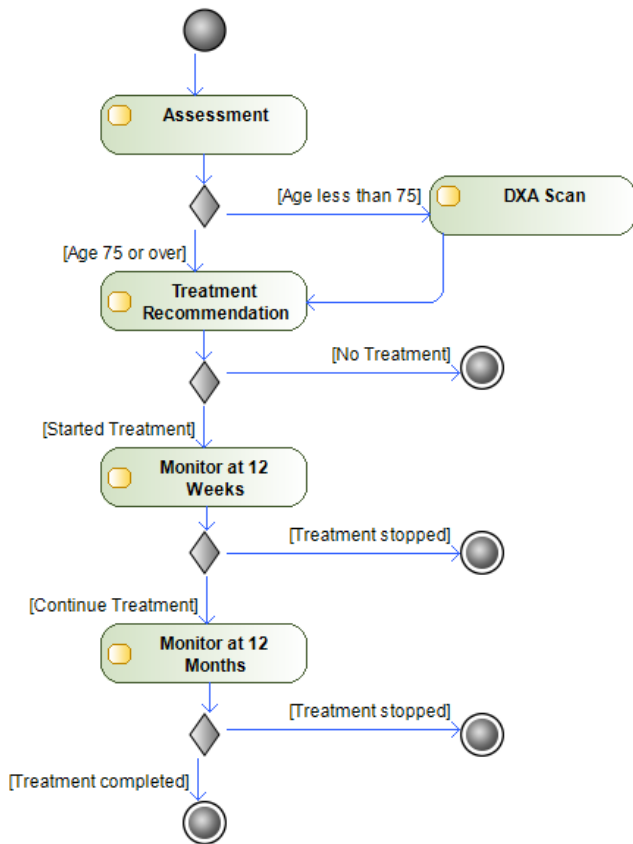

Figure 1- Simple Care Pathway for Fracture Liaison Service.

The pathway starts when patients are identified by the service, following hospital admission with a diagnosed fragility fracture (a fracture caused by a fall or minor trauma that would not normally result in a fracture, unless the bones had been weakened). The identified patient is assessed and sent for a bone density scan (DXA) if they are under the age of 75 (patients aged 75 and over are assumed to have some degree of age-related bone weakness, so the bone density scan is not deemed necessary). In most cases, a recommendation is made for treatment, but if not, the pathway ends. Patients who do have treatment are monitored after 12 weeks and then again at 12 months, to check compliance to the treatment, recommend possible alternatives or to acknowledge the patient's decision to stop treatment.

The cityEHR stores all clinical data as HL7 CDA documents, which follow the hierarchical structure of a health record specified in the ISO-13606 standard [7]; folder, composition, section, entry and element, where the elements record the actual recorded data values and the entries that contain elements are the lowest level of atomic clinical context.

All clinical data gathered in the example pathway shown in Figure 1, and any clinical decisions made, are recorded in the patient record in CDA documents for the Assessment, DXA Scan Results, Treatment Recommendation and Monitoring. The Monitoring documents are completed based on the answers to a questionnaire sent to patients in a letter, prior to the 12 week and 12 month monitoring dates; these letters are also stored in the patient record as CDA documents.

A formal specification of the pathway is created in cityEHR as a CDA document, which can then be instantiated for a specific patient and saved in the record in a series of versions that record progress through the sequence of actions in the pathway and prompt users (through their In-Tray) for the next action(s) to perform [3]. Once the pathway is completed, the CDA document is committed to the patient record as evidence of the actions performed.

Informal Implementation of Care Pathways

The same core features of cityEHR that enable the formal execution of care pathways, as described above, can be used in a more loosely coupled way to achieve a similar control over the sequence of clinical activities. These features are:

[1] support for care teams;

[2] triggers and notification when documents are committed to the record;

[3] annotation of the historic record, with associated notifications;

[4] a clinician-centred In-Tray for handling of notifications.

Support for Care Teams

The cityEHR supports the formation of care teams, which are composed of both clinical and administrative system users. Subject to access control permissions, any user can create a care team and then assign users to it. The care teams act as user groups for the purposes of clinical communication. Any

such communication (initiated through the notification or annotation mechanisms described below) can be directed to an individual (named) user, to a user role or to a care team.

Triggers and Notifications

Triggers can be configured to watch for the committal of specific compositions, entries and elements to the record. There are two types of trigger:

- for immediate navigation/action by the current user

- to notify future actions that may be performed by any (specified) user

Triggers are defined so that when a specific combination of composition, entry and element is committed to the record, and other pre-set conditions are satisfied, the trigger is activated.

For navigation triggers, the user is immediately prompted for the choice of navigation to other parts of the patient record. These choices can include viewing the clinical dashboard, viewing summaries of the record for the individual patient or creation of a new clinical document for the patient; for example a form, letter, order or prescription.

For notification triggers, a new notification composition is instantiated and added to the patient record. Notifications are stored as HL7 CDA documents, in the same way as all other clinical data, and contain information about the source of the notification, the system users to be notified and the actions (if any) that are required in response.

Annotation of the Historic Record

The mechanism of triggers and notification described above, implements pre-configured pathways of clinical actions. For more ad-hoc communication between members of the care team, the annotation features can be used. Once data have been committed to the record (in the form of an HL7 CDA composition), users can annotate any document in the record and address notifications to other users, roles of user or members of a care team. It is not possible to send ad-hoc notifications without reference to a document in the record. This ensures that all communication is centred around the patient record; there is no support for communication between system users that is not related to an individual patient.

Clinician-Centred In-Tray

The In-Tray acts as the hub for communication, whereby a system user can view and act upon all the notifications that have been addressed to them by other users. These notifications may have been addressed to the individual user, to one of their assigned user roles or to one of the care teams of which they are a member.

The In-Tray is a cross-patient feature, which looks and feels similar to a standard email facility. However, it is fundamentally different, in the sense that the list of notifications seen in a user's In-Tray is formed by a dynamic query across the records stored in the cityEHR. Each notification then provides a link directly through to the record of the individual patient about which it concerns. Hence, there is no stored 'in-box' of notifications for each user, instead there is a display of the virtual in-box created each time the user accesses the In-Tray. Similarly, although the user accesses specific documents, or actions, in the patient record through the In-Tray, no part of the patient record is transmitted; the notification merely contains a pointer to the patient record, which remains stored in the database and subject to the usual access controls.

Results

Reporting of Variance from Formal Pathway

Using the methods described in the previous section, a care pathway can be implemented using informal mechanisms to guide the actions of system users. As these actions are completed, the EHR provides a history of the results of each action in the form of HL7 CDA documents stored in the record for the patient. To the system user, this process seems similar to the more formal implementation, but unlike the formal implementation there is no underlying 'engine' to ensure that the sequence of actions follows the pre-defined pathway.

Instead the user is guided by a series of notifications or annotations, each of which is triggered independently, based on the current state of the patient record. As a result, since no predetermined sequence of actions is being followed, it is not possible to determine at any specific time, whether a particular pathway has been completed, or whether a pathway is in progress and partially completed.

However, we may wish to analyse the sequence of actions completed, to assess whether they follow the steps in a formally specified pathway which has been applied retrospectively.

To achieve this, the pathway is defined as an HL7 CDA document, in the same way that it would be defined to drive the formal execution of the pathway. The CDA document (in XML) is then transformed using XSLT into a representation of the finite states in the pathway. This representation uses the State Chart Extensible Markup Language (SCXML) which was published as a W3C standard in 2015 [10]. SCXML was based, in part, on the STATEMATE graphical process representation language, first described by Harel et al [11].

Using the SCXML representation, a set of database queries is generated to interrogate the state of the patient record along a set of time lines, derived from the set of possible initial states found in the record. The queries are expressed in XQuery, the W3C standard for representing queries for XML databases, since the EHR is stored as a collection of HL7 CDA (XML) documents.

Hence, the first query finds any documents recorded in the patient record which match the initial state for the pathway. Each of these documents is a candidate initial state and the effective time of the document (i.e. the time it was committed to the record) sets the base time for the pathway. The next queries find candidates for the second state in the path-

way and advance the timeline to the effective time of those states.

The sequence of XQuery submissions to the database is continued until a sequence of states has been found which validates the pathway. Alternatively, the submission of searches continues until all the set of possible valid sequences has been exhausted. The results for the searches are then collated and returned as the variance report, which is generated as an HL7 CDA document (which can then be stored in the patient record, if required).

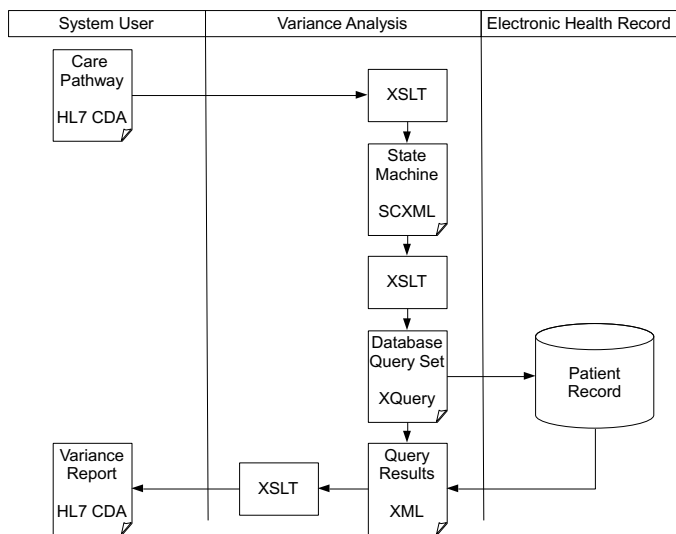The overall method of reporting variance is summarised in Figure 2.

Figure 2-Reporting Variance Using SCXML.

Returning to the example pathway from Figure 1, it can be seen that there is a single initial node in the pathway and four possible end nodes. The state transition from the initial node to one of the end nodes can be defined by a sequence of CDA documents, as shown in Figure 3.

The state transitions are shown for patients under the age of 75 years, who have a DXA scan; there is a similar set of transitions for patients aged 75 and over, without the DXA Results document. The set of XQueries generated from the pathway definition search for patterns of documents that match the state transitions. So in this example, the first query searches for Assessment documents, then DXA Results, then Treatment Recommendations, in the correct time order. Within the documents, the queries test for Age < 75 and the type of Treatment (Started or None). The next queries search for Monitoring documents, with tests for the Monitoring Phase (12 weeks or 12 months) and the Treatment (Ongoing, Stopped or Completed).

Collating the results of these queries allows a report of variance to be generated, which shows:

- potential pathways that were started but then deviated from the definition (had one or more valid transitions from the start state, but then had an invalid transition)

- pathways that have started, with valid transitions, but which are not complete (are missing states to the end node, but have no invalid transitions)

- pathways which are complete (have a set of valid transitions from a start node to and end node)
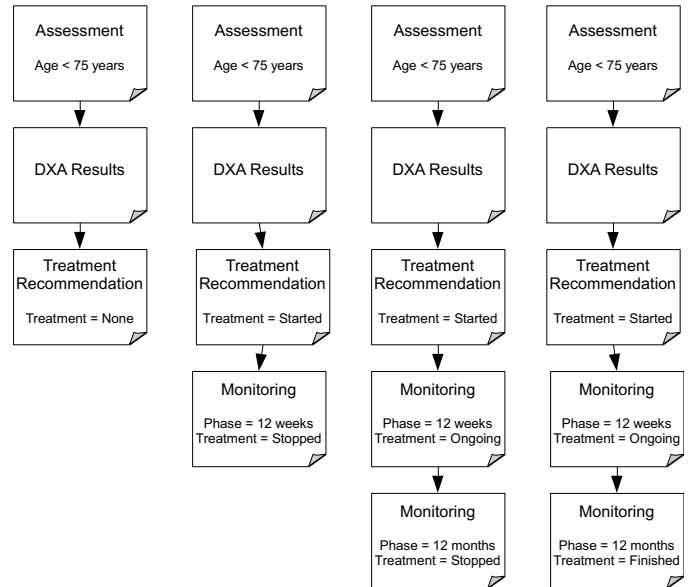

Figure 3: Pathway State Transition

Discussion and Conclusion

Care pathways in cityEHR are modelled as HL7 CDA documents. The XML structure of an HL7 CDA document is ideal for representing the hierarchical decomposition of tasks in a pathway, and processing of the pathway can take advantage of some general properties of XML as a hierarchy of document nodes. We have demonstrated how the same methodology can be used in the implementation of informal care pathways. However, by describing care pathways using informal mechanisms it becomes a requirement to determine the completion of a particular pathway according to the actions recorded in the patient record.

We propose a method to report on the variance from a formally defined pathway, using the information set recorded for the patient. This method uses the same representation of the pathway that would be used in the formal implementation but, transforms it first into a finite state machine representation (SCXML), and then into a set of (XQuery) database queries that are used to find patterns of information in the record and report on variance of the recorded data from the pathway.

One key advantage of the method we propose is that any pathway can be formulated retrospectively, and then applied against the patient record to report variance.

The implementation of this method has been validated with simple example pathways, such as the one presented in this paper, and has been shown to function correctly. Future work

will be to validate the method with more complex pathways and real patient data.

References

[1] Gooch P, Roudsari A. Computerization of workflows, guidelines, and care pathways: a review of implementation challenges for process-oriented health information systems. Journal of the American Medical Informatics Association. 2011;18(6):738-748.

[2] Dolin RH, Alschuler L, Boyer S, Beebe C, Behlen FM, Biron PV, et al. HL7 clinical document architecture, release 2. Journal of the American Medical Informatics Association. 2006;13(1):30-39.

[3] Chelsom J, Pande I, Gaywood I, Granja C, Bolle SR. Document-Driven Care Pathways Using HL7 CDA. In: Proceedings of the Seventh International Conference on eHealth T, and Social Medicine. International Academy, Research and Industry Association; 2015 Presented at: eTelemed February 22-27, 2015; Lisbon, Portugal p. 118-123.

[4] Ueckert F, Goerz M, Ataian M, Tessmann S, Prokosch H-U. Empowerment of patients and communication with health care professionals through an electronic health record. International journal of medical informatics. 2003;70(2-3):99-108.

[5] White A, Danis M. Enhancing patient-centered communication and collaboration by using the electronic health record in the examination room. Jama. 2013;309(22):2327-2328.

[6] Bitton A. Who is on the home team? Redefining the relationship between primary and specialty care in the patient-centered medical home. LWW; 2011.

[7] Wakamiya S, Yamauchi K. What are the standard functions of electronic clinical pathways? International journal of medical informatics. 2009;78(8):543-550.

[8] Eshuis R. Symbolic model checking of UML activity diagrams. ACM Transactions on Software Engineering and Methodology (TOSEM). 2006;15(1):1-38.

[9] Awad A, Decker G, Weske M. Efficient compliance checking using BPMN-Q and temporal logic. In: Proceedings of the Springer; Presented at: International Conference on Business Process Management, 2008; p. 326-341.

[10] Barnett J, Akolkar R, Auburn R, Bodell M, Burnett D, Carter J, et al. State chart XML (SCXML) state machine notation for control abstraction. W3C Recommendation. 2015.

[11] Harel D, Lachover H, Naamad A, Pnueli A, Politi M, Sherman R, et al. Statemate: A working environment for the development of complex reactive systems. IEEE Transactions on software engineering. 1990;16(4):403-414.

Address for correspondence

Dr John Chelsom, Seven Informatics, 16a North Bar Street, Banbury, Oxon, UK. john.chelsom@seveninformatics.com

Which Factors of Business Intelligence Affect Individual Impact in Public Healthcare?

Rikke GAardboe[a], Niels Sandalgaard[a] , Tanja Svarre [b]

[a]*Department of Business and Management, Aalborg University, Aalborg, Denmark*
[b] *Department of Communication and Psychology, Aalborg University, Aalborg, Denmark*

Abstract

In this paper, we examine the relationship between business intelligence (BI) quality, task characteristics and individual impact of the system from an end-user perspective at 12 public hospitals. 1,352 BI end-users answered the questionnaire. Linear regression was used to test the research model empirically. If organisations in the public health sector want high individual impact, the following factors are essential. Firstly, system quality must be high. Secondly, the system must support the tasks that the BI user solves with the system. Thirdly, task difficulty is positively and significantly related to impact. In conclusion, it is essential that the user perceives the task as being important. The user's perception of task interdependency and task specificity does not influence individual impact. Future research should focus on different healthcare settings with different types of BI system.

Keywords:

Business intelligence, Public healthcare, End-user success.

Introduction

Public healthcare sectors generate large amounts of data relating to patient records, compliance, and patient care [1]. Therefore, there is an increasing interest in using business intelligence (BI) within both the private and public health sectors. 'BI' is "…commonly used to describe the technologies, applications, and processes for gathering, storing, accessing, and analysing data to help users make better decisions" [2]. This definition implies that if BI is utilised to enhance decision-making, it can affect an organisation's performance. According to Mettler and Virmarlund, the value of BI in a healthcare setting is not only in providing information but also that "…its contribution is in enabling new ways of working, allowing the integration of information and organisations and the measurement of outputs in real time" [3].

One issue that has dominated research in information systems (IS) is how organisations can achieve success. In the literature, there are countless definitions and goals for success [4]. In this article, success will be measured as the dependent variable "individual impact", because there is a relationship between the individual impact and the organisational impact of IS [5]. In this study, success is based on an end-user perspective. Success in public healthcare is interesting because there is a significant amount of data and a complex system landscape [6]. There is a difference between evaluating IS in the public and private sectors [7], and most studies have been conducted in private organisations [4, 8] while public sector studies are lacking. In this regard, our goal is to assess the individual impact of business intelligence in public hospitals and associated administrative areas in Denmark. Unlike many other organisations, there is a high degree of diversity in BI users in healthcare, since BI is used to support both administrative and clinical decisions. Therefore, user types range from administrative employees to clinical staff and they solve many different types of task with BI [9]. Hence, it is interesting to look at the relationship between task characteristics, BI quality, and individual impact in particular in a healthcare setting[10].

In this paper, we test the relationship between five task characteristics, information quality, system quality and individual impact in 12 public hospitals and their administration. The article contributes to the subfield of the relationship between task and technology in public healthcare. The article is organised as follows: in the next, second section, we present related literature and the research model, and discuss our methods in the third part In the fourth section, we analyse the findings, which are then discussed. The final section covers our conclusions.

Research model

In recent years, there has been an increasing amount of literature on the factors for BI success [11]. Traditionally, these studies have been based on one or several of the success measures from DeLone and McLean's IS success model [5, 12]. One measure of BI success is individual impact [4], and DeLone and McLean defined this as "an indication that an information system has given the user a better understanding of the decision context, has improved his or her decision-making productivity, has produced a change in user activity, or has changed the decision maker's perception of the importance or usefulness of the information system" [5].

Numerous independent variables have been investigated for their effect on individual impact. However, few studies have examined the relationship between task characteristics and individual impact from an end-user perspective [11, 13]. Tasks are activities that support an organisation, and the role of IT systems is to support the completion of these [14]. The purpose of using IS systems is to inform or to automate tasks [15]. Petter et al. [13] have categorised task characteristics in relation to information systems to include task compatibility, task significance, task difficulty, task interdependence and, moreover, task specificity. The fit between the user's task and BI is referred to as 'task compatibility' [13]. 'Task interdependence' reflects whether the completion of a BI-related task depends on others, while the importance of the task is characterised as 'task significance' [13]. 'Task difficulty' is the extent to which a user believes that BI makes it possible to complete complicated tasks [13]. Finally, 'task specificity' is the level of clarity of the task supported by BI [13].

BI quality can be assessed using system quality and information quality [16]. Ease of use, data quality, and maintenance of the BI system are referred to as 'system quality' [17], while the quality of the output from BI is known as 'information quality' [17]. As use is related to the system's yield, it is measured regarding the time needed for use [17]. The use of information systems for specific tasks is often mandated, therefore user satisfaction is measured in relation to the particular system [18]. In this study, the user's overall satisfaction with BI is measured.

The central question in this article is how task characteristics and BI quality are related to individual impact. In previous research, single tasks have been tested against individual impact, but no studies have tested all of the above task characteristics with individual impact [13] The research model is presented end in the figure below.
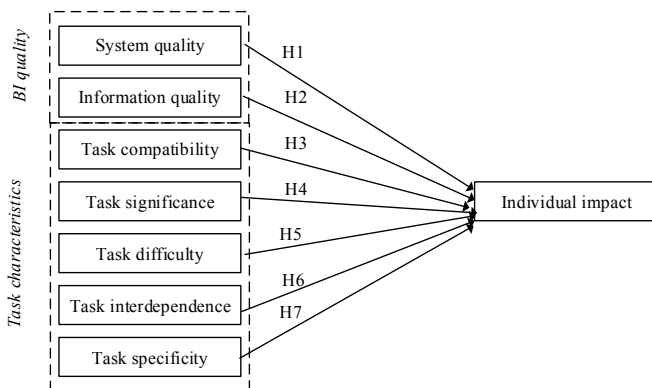

Figure 1- The research model

Based on the above, the following hypotheses can be proposed:

H1: There is a positive relationship between system quality and individual impact

H2: There is a positive relationship between information quality and individual impact

H3: There is a positive relationship between task compatibility and individual impact

H4: There is a positive relationship between task significance and individual impact

H5: There is a positive relationship between task difficulty and individual impact

H6: There is a positive relationship between task interdependence and individual impact

H7: There is a positive relationship between task specificity and individual impact

Methods

Data collection process

To test the hypotheses, 12 public hospitals in the same region in Denmark were selected as research sites. All employees who had access rights to BI were chosen as the sample, including members of management, nurses, doctors, and administrative staff, etc. First, an email was sent out explaining the project and encouraging employees to participate. Afterwards, each respondent received an email with a link to the questionnaire, which was prepared using an online survey program. The first question focused on whether the respondent had used the BI system. If the respondent answered "No", no further questions were asked. Later, a reminder was sent two weeks after the initial distribution of the survey to encourage respondents to complete the questionnaire. Before the questionnaire was sent out, questions were selected by a comprehensive literature review [19] and a test among senior researchers in the relevant field of research. Of the 4,232 employees invited to participate, 1,351 responded, giving a response rate of 32%. Of these, 605 indicated that they did not use the BI system. This left 746 responses to be used for the data analysis.

Questions used in the survey

Our questions, shown in Table 1, have all been used in previous information systems studies [5, 12, 20–22]. However, we translated the questions into Danish and ensured that they fitted the specific context of our sample. The survey was part of a larger research project and so, in addition to the questions listed in Table 1, the questionnaire contained questions requesting data not used in this particular paper [6, 10]. In the questionnaire [BI] was replaced with each organisation's term for BI, for instance 'Tableau'.

Table 1- Constructs, questions and Cronbach's alpha

Construct	Question	Cronbach Alpha
Individual Impact	I can efficiently make my reports using BI. [20]	0.844
	I can complete my reports quickly using BI. [20]	

Construct	Question	Cronbach Alpha
	I can complete my reports using BI. [20]	
System Quality	BI is easy to learn.[20]	0.826
	BI is easy to use. [21]	
	The information in BI is easy to understand.[22]	
Information Quality	Data are displayed in a consistent format in BI. [22]	0.744
	The data in BI have high validity. [22]	
	Other employees in the region also think the data in BI have high validity. [22]	
Task compatibility	This information is useful for my tasks.[22]	0.817
	This information is complete for my needs. [22]	
	This information is relevant to my tasks. [22]	
	This information is relevant to my tasks. [22]	
Task significance	The tasks I complete in BI are an important part of my tasks. [23]	0.776
	I make decisions based on the tasks I complete in BI. [23]	
	My tasks completed in BI are important to other employees in the organisation. [23]	
	Other people make decisions based on the tasks I completed in BI. [23]	
	My tasks in BI are important for collaborators outside the organisation. [23]	
Task difficulty	BI makes it possible to complete complicated tasks. [23]	-
Task Interdependence	If I do not complete my tasks in BI, one or more employees in the organisation cannot complete their tasks. [23]	-
Task Specificity	The tasks I complete in BI can be done in more than one way.[24]	-

Cronbach's alpha is calculated in the third column of Table 1 for the following constructs: individual impact, system quality, information quality, task compatibility and task significance. All values are above the threshold value of 0.7 [25], which indicates good reliability. Task difficulty, task interdependence, and task specificity represent single items and Cronbach's alpha is not therefore calculated.

Findings

The hypotheses are tested using multiple linear regression. Zviran et al. [26] have put forward a hypothesis regarding gender and age being antecedents of individual impact. Therefore, we have chosen to include these as control variables. Consequently, the following regression was used:

$$Y=a+b_1x_1+b_2x_2+b_3x_3+b_4x_4+b_5x_5+b_6x_6+b_7x_7+b_8x_8+b_9x_9+e$$

where Y is individual impact, X_1 is system quality, X_2 is information quality, X_3 is task compatibility, X_4 is task significance, X_5 is task difficulty, X_6 is task interdependence, X_7 is task specificity, X_8 is gender, X_9 is age, and finally, e is the error term.

Table 2- Regression results

Y=individual impact	B	Std. Error	Beta	t	Sig.
(Constant)	-,207	,180		-1,153	,250
System quality	,505	,044	,440	11,567	,000
Information quality	,089	,054	,069	1,646	,100
Task compatibility	,229	,051	,187	4,466	,000
Task significance	,145	,040	,129	3,586	,000
Task difficulty	,098	,037	,098	2,664	,008
Task interdependence	-,021	,026	-,028	-,813	,417
Task specificity	-,030	,035	-,028	-,864	,388
Gender	,002	,004	,015	,495	,621
Age	-,010	,007	-,044	-1,434	,152
Notes: R^2=0,463 Adj. R^2=0,455					

The results of the regression are shown in Table 2. When inspecting the main effects, it can be seen that system quality and task compatibility are positively related to individual impact ($p<0.001$). Task difficulty is also positively and significantly associated with individual impact ($p<0.01$). It can also be seen that information quality, task interdependence and task specificity are insignificant. The two control variables (gender and age) are also insignificant. In sum, system quality, task compatibility, task significance and task diffi-

culty have an impact on whether the user experiences individual impact through using BI.

Discussion

Based on the regression analysis, the findings of this study regarding the seven hypotheses are shown in Table 2.

Higher system quality leads to higher individual impact from employees in the public health sector using BI (H1). If consideration is given to the context in which BI is evaluated, it can be concluded that employees in the public health sector will experience higher individual impact if the system is easy to understand, learn and use.

Surprisingly, there is an insignificant relationship between information quality and individual impact (H2). Therefore, higher information quality does not necessarily lead to higher individual impact of BI. The finding is surprising because other studies, for instance, D'Ambra and Rice [27] and Shih [28] find a relationship between information quality and both the quality of work and time savings. One reason why the relationship between information quality and individual impact (H2) is insignificant may be that users will not experience task compatibility if there is inadequate quality of information. Therefore, information quality can be perceived as a hygiene factor that is assumed to be present.

Access is the relationship between task compatibility and individual impact, which was found to be positive and significant (H3). That is, if an employee's tasks and the system have a positive fit, they will find that the tasks can be completed quickly and efficiently. Several studies have confirmed this positive and significant relationship [29, 30].

Two other task characteristics that are positive and significantly related to individual impact are task significance and task difficulty (H4 and H5). The more critical the user considers the tasks that are solved with BI, the higher the individual impact (H4). Few other studies have confirmed the relationship, even though the finding was expected [13]. One explanation may be that by completing complicated tasks quickly and efficiently, the user perceives the individual impact of BI to be higher. The same relationship applies to task difficulty; the more difficult a task is for the BI user to solve, the higher they consider the individual impact of using BI (H5). A reason for this may be that they perceive BI as a means of solving their tasks.

The relationship between task interdependence and individual impact (H6) as well as the relationship between task specificity and individual impact (H7) were insignificant. In the literature, there is mixed support for this particular relationship; a study by Kim et al. [31] discovers support for it, while Marchal et al. [32] does not.

In summary, there is a relationship between system quality, task compatibility, task significance, task difficulty and the users' individual impact of BI in the public sector. The remainder of the relationships tested are insignificant. These findings are supported in the literature, but no studies have previously been performed with so many task characteristics in a BI and public sector setting to identify which factors contribute to individual impact.

Conclusion

The primary goal of the current study was to determine which BI quality and task characteristics lead to individual impact in a public healthcare setting. This is a contribution, as this has not been widely researched before in a healthcare setting. If organisations in the public health sector want high individual impact for BI users, the following things are essential. Firstly, system quality must be high. Secondly, the system must support the tasks that the BI user solves with it. In addition, task difficulty is also positively related to impact. In conclusion, it is essential that the user perceives the task as being important.

The findings in this article are subject to more limitations. First, the study is conducted in one specific country and only in the public sector. Moreover, only one specific type of BI system has been investigated. Therefore, future research should focus on different healthcare settings with different types of BI systems and take national differences into account. Also, the operationalisation of individual impact could be further developed. We use the construct by Wang & Liao[21] who focus on the effect the BI system has on the ability to make reports but other ways of measuring impact could also be tested. Likewise, some of the measures focus on the perception of the user. Further research should consider if some of these measures could be measured more directly.

References

[1] Raghupathi W, Raghupathi V. An overview of health analytics. J Health Med Informat 2013; 4: 2.

[2] Wixom B, Watson H. The BI-Based Organization: International Journal of Business Intelligence Research 2010; 1: 13–28.

[3] Mettler T, Vimarlund V. Understanding business intelligence in the context of healthcare. Health informatics journal 2009; 15: 254–264.

[4] Tona O, Carlsson SA, Eom S. An empirical test of Delone and McLean's information system success model in a public organization. In: 18th Americas Conference on Information Systems 2012, AMCIS 2012. 2012, pp. 1374–1382.

[5] DeLone WH, McLean ER. Information Systems Success: The Quest for the Dependent Variable. Information Systems Research 1992; 3: 60–95.

[6] Gaardboe R, Sandalgaard N, Nyvang T. An assessment of business intelligence in public hospitals. IJISPM - International Journal of Information Systems and Project Management 2017; 5–18.

[7] Rosacker KM, Olson DL. Public sector information system critical success factors. Transforming Government: People, Process and Policy 2008; 2: 60–70.

[8] Kolasa I. Success Factors for Public Sector Information System Projects: Qualitative Literature Review. 2017, p. 326.

[9] Madsen L. Healthcare business intelligence: a guide to empowering successful data reporting and analytics. Hoboken, New Jersey: John Wiley & Sons, Inc., http://catalogimages.wiley.com/images/db/jimages/97811 18217801.jpg (2012, accessed 31 July 2017).

[10] Gaardboe R, Sandalgaard N, Sudzina F. The importance of task compatibility for web-enabled Business Intelligence success in e-government. In: Proceedings of the 19th International Conference on Information Integration and Web-based Applications & Services (iiWAS2017). Salzburg: International Conference on Information Integration and Web-based Applications & Services, 2017.

[11] Gaardboe R, Svarre T. Business Intelligence Success Factors: A literature review. Journal of Information Technology Management 2018; 29: 1–15.

[12] DeLone WH, McLean ER. The DeLone and McLean Model of Information Systems Success: A Ten-Year Update. Journal of Management Information Systems 2003; 19: 9–30.

[13] Petter S, DeLone W, McLean ER. Information Systems Success: The Quest for the Independent Variables. Journal of Management Information Systems 2013; 29: 7–62.

[14] Leavitt HJ. Applied organizational change in industry: Structural, technological and humanistic approaches. In: March J (ed) Handbook of Organizations. 1965, pp. 1144–1170.

[15] Zuboff S. In the age of the smart machine: the future of work and power. Oxford: Heinemann, 1988.

[16] Gaardboe R, Svarre T, Nyvang T. The Relationship between Task Characteristics, BI Quality, and Task Compatibility: An Explorative Study. Complex Systems Informatics and Modeling Quarterly 2018; 14: 54–63.

[17] Seddon PB. A Respecification and Extension of the DeLone and McLean Model of IS Success. Information Systems Research 1997; 8: 240–253.

[18] Hsieh J, Rai A, Petter S, et al. Impact of User Satisfaction with Mandated CRM Use on Employee Service Quality. MIS Quarterly 2012; 36: 1065.

[19] Gaardboe R, Svarre T. Critical Success factors for Business Intelligence Success. 2017, pp. 472–486.

[20] Lewis JR. IBM Computer Usability Satisfaction Questionnaires: Psychometric Evaluation and Instructions for Use. International Journal of Human-Computer Interaction 1995; 7: 57–78.

[21] Wang Y-S, Liao Y-W. Assessing eGovernment systems success: A validation of the DeLone and McLean model of information systems success. Government Information Quarterly 2008; 25: 717–733.

[22] Lee YW, Strong DM, Kahn BK, et al. AIMQ: a methodology for information quality assessment. Information & management 2002; 40: 133–146.

[23] Morgeson FP, Humphrey SE. The Work Design Questionnaire (WDQ): Developing and validating a comprehensive measure for assessing job design and the nature of work. Journal of Applied Psychology 2006; 91: 1321–1339.

[24] Daft RL, Macintosh NB. A Tentative Exploration into the Amount and Equivocality of Information Processing in Organizational Work Units. Administrative Science Quarterly 1981; 26: 207–224.

[25] Nunnally JC, Bernstein IH. The assessment of reliability. Psychometric theory 1994; 3: 248–292.

[26] Zviran M, Pliskin N, Levin R. Measuring user satisfaction and perceived usefulness in the ERP context. Journal of Computer Information Systems 2005; 45: 43–52.

[27] D'Ambra J, Rice RE. Emerging factors in user evaluation of the World Wide Web. Information & Management 2001; 38: 373–384.

[28] Shih H-P. Extended technology acceptance model of Internet utilization behavior. Information & management 2004; 41: 719–729.

[29] Goodhue DL, Thompson RL. Task-Technology Fit and Individual Performance. MIS Quarterly 1995; 19: 213–236.

[30] Jarupathirun S, Zahedi F "Mariam". Dialectic decision support systems: System design and empirical evaluation. Decision Support Systems 2007; 43: 1553–1570.

[31] Kim DT, Kim BG, Aiken MW, et al. The influence of individual, task organizational support, and subject norm factors on the adoption of Groupware. Academy of Information and Management Sciences Journal 2006; 9: 93–110.

[32] Marshall TE, Byrd TA, Gardiner LR, et al. Technology acceptance and performance: An investigation into requisite knowledge. Information Resources Management Journal 2000; 13: 33–45.

Address for correspondence

Rikke Gaardboe, Department of Business and Management, Fibigerstrade 11, 9220 Aalborg Oest, Denmark.

Email: gaardboe@business.aau.dk

Usability Evaluation of a Smartwatch Heart Rate Monitor for Subjects with Acquired Brain Injury

Morten Pallisgaard Støve , Birgit Tine Larsen

Department of Physiotherapy, University College of Northern Denmark.

Introduction

Cardiorespiratory exercise at specific heart rate reserve intensity levels is recommended in the national Danish clinical guideline for rehabilitation of acquired brain injury [1].

Traditional pulse rate monitors utilizing chest-straps are used in clinical practice, but they cannot be worn for extended periods [2]. They may also be inconvenient to use [3] and hemiplegic patients may require therapeutic assistance to apply the chest-strap.

New consumer-based wrist-worn activity monitors that utilize photoplethysmography technology to measure heart rate demonstrate good potential to overcome the limitations of the traditional chest-strap monitors in rehabilitation settings [4]. However, the usability for patients with acquired brain injury is currently unknown, hence the purpose of this study was to examine the usability of heart rate measurements using a smartwatch for patients with acquired brain injury.

Materials and Methods

Twenty-five subjects with acquired brain injury were enrolled in the second week following admittance to an inpatient rehabilitation clinic. Subjects wore a smartwatch (Garmin Forerunner 235) during daily rehabilitation activities for a period of three weeks.

Two Likert scale-based questionnaires were used to elicit both the subject's assessment of their own ability to handle the watch and the nature and severity of difficulty the subject demonstrated while handling the watch evaluated by the physio- or occupational therapists. Field observations, recorded by the therapists as field notes using a standardised observation log, were used to record the interactions between the subjects and the smartwatch and to identify key usability problems during rehabilitation activities.

Normality of the data was assessed by the Shapiro-Wilks test. Participants were ranked into three groups according to Functional Independence Measure (FIM) scores to make comparisons between groups. Participant and therapist ratings were analysed using descriptive statistical methods and the statistical difference between groups of different items was tested using Wilcoxon signed rank sum tests. The findings from the observations were summarized into themes using an inductive thematic approach.

Results

Quantitative data showed that most subjects were able to put on the watch, read the pulse indicator and charge the watch independently, however, only a few subjects were able to start and stop pulse recordings independently.

Qualitative observations revealed; 1) That only a few subjects operated the watch at their own initiative, hence most of the subjects had to be reminded by the therapist to do so, daily, during the three-week period. 2) Subjects were more often baffled by the many functionalities of the watch and generally found the interface unintuitive and confusing.

Subjects with severe motor and cognitive impairment had significantly more difficulty in putting on the watch compared with subjects with less impairment ($p < 0.039$). There was no difference between the subject's own assessment of their ability to handle the watch and the therapeutic assessments of the subject's ability to handle the watch ($p > 0.555$).

Discussion

Based on the results of the present study the usability of the Garmin Forerunner 235 for patients with acquired brain injury proved relatively low as most participants were unable to independently operate the watch during the three-week period.

The main finding was that most subjects, although physically capable, did not independently operate the watch during rehabilitation activities. This may be attributed to the nature and severity of the injuries and a reduced ability to develop automated responses which may have been exacerbated by the unintuitive watch interface.

It should, however, be noted, that n = 17 subjects had moderate to severe impairments of memory, understanding and problem-solving domains of the cognitive FIM which could affect the results of this study.

References

[1] Sundhedsstyrelsen (2014) National klinisk retningslinje for fysioterapi og ergoterapi til voksne med nedsat funktionsevne som følge af erhvervet hjerneskade, herunder apopleksi : 8 udvalgte indsatser

[2] Fallow AB, Tarumi T, Tanaka H (2013) Influence of skin type and wavelength on light wave reflectance. J Clin Monit Comput 313–317 .

[3] Spierer DK, Rossen Z, Litman LL, Fuji Kenji (2015) Validation of photoplethysmography as a method to detect heart rate during rest and exercise. J Med Eng Technol 264–271 .

[4] Hickey AM, Freedson PS (2016) Utility of Consumer Physical Activity Trackers as an Intervention Tool in Cardiovascular Disease Prevention and Treatment. Prog Cardiovasc Dis 58:613–619 .

Address for correspondence

Morten Pallisgaard Støve: mps@ucn.dk

Recruitment to and Dropouts from Telemedicine Interventions

Carl Erik Moe[a] , Elin Thygesen[b]

[a]*Department of Information Systems, University of Agder, Norway*
[b]*Department of Health and Nursing Science, University of Agder, Norway*

Introduction

Telemedicine interventions are increasingly applied to cater for patients living at home with chronic diseases. Studies show promising results, but still more research is needed. In several of these interventions recruitment of patients is difficult, and the attrition rate is high. The telemedicine project TELMA (Telemedicine Agder) is a large R&D (research and development) project with a goal of recruiting 700 patients, with COPD, diabetes, heart congestion and multimorbid patients.

The TELMA is run by three municipalities in Agder, Norway, in collaboration with the regional hospital trust, University of Agder, and two vendors. The goal of the project is to establish a service for all the 30 municipalities in the region, for groups of patients that yield clear benefits in form of less use of health personnel, and at the same time equal or increased life quality. Patients are recruited from different entities, hospital, the municipalities and GPs. As this is a research project they all give informed consent to researchers, and accept answering a survey two or more times, being interviewed and granting the researchers access to different datasets.

As the municipalities carry the expenses of the services, the municipalities require a formal decision regarding granting the service to the patients. The patients are granted the service for a period of 3 months, after which a new decision is made whether to carry on the service or not, based on the outcome so far. The service is offered to the 30 municipalities from one of three telemedical centrals. The centrals are staffed by a nurse, and a GP can be consulted if needed. The nurses exchange electronic messages with the patients´ GP regularly. The service consists of daily monitoring of patient data and a self-assessment survey, and regularly video conference with the nurse.

However; recruitment has taken longer time than expected, and through our study we are identifying possible reasons for declining to participate and for dropping out. Hence; our research questions are: "Are the "right patients" recruited, and why do patients drop-out from this intervention?

We know from previous studies [1, 2] that recruitment of patients and drop-outs is a problem in many health interventions and believe that our results will of general interest.

Materials and Methods

An interpretive approach was chosen, interviewing patients and different stakeholders involved in recruitment and follow-up.

The employees are recruited through contacting the assistant project manager. The employees include a nurse from the hospital, 4 nurses from 2 different telemedical centrals, one project participant involved in recruitment, one nurse employed at a GP´s office, and one case handler from the health care services. The patients are recruited from the telemedical centrals. This group includes 5 patients that have participated for a period up to 3 months and 3 that have dropped out. We have asked and will ask the patients about their condition and their needs and living conditions (whether they live alone or have a next of kin living with them), about how they were recruited, and about the introduction to the service. We also ask them about the service itself, and the reason for dropping out for the ones that drop-out. The employees are asked about the recruitment process and the introduction to the service.

As some of the patients are severely ill, we have been careful not to tire them unnecessary, and have made it clear that we are ready to stop the interview if wanted. All participants have given formal consent to participate in the interviews. The project has been approved by the Norwegian Center for Research Data (project number 53693).

Results and Discussion

We have only preliminary data, and are ready to present and discuss the results at the conference, and these will also carry implications for the project

References

[1] Eysenbach, G. (2005). The law of attrition. *Journal of medical Internet research*, *7*(1).

[2] Kannisto, K. A., Korhonen, J., Adams, C. E., Koivunen, M. H., Vahlberg, T., & Välimäki, M. A. (2017). Factors associated with dropout during recruitment and follow-up periods of a mHealth-based randomized controlled trial for Mobile. Net to encourage treatment adherence for people with serious mental health problems. *Journal of medical Internet research*, *19*(2).

Address for correspondence

Carl.E.Moe@uia.no and Elin.Thygesen@uia.no

Exploring the Benefits and Challenges of Tele-Healthcare: A Multiple Case Study of the Use of individual Video-Consultations in Alcohol Addiction Undergoing Withdrawal Treatment and Sexual Counselling in Denmark

Ulla Virkkunen Andrees[a], Bo Bojesen[b], Karsten Ulrik Niss[c]

[a]*Midwife, Hospital of Southern Jutland, Region of Southern Denmark, Denmark*
[b]*MD, Specialist in Psychiatry, Psychiatric Treatment Facility, Capital Region, Denmark*
[c]*Research Assistant, Department of Business and* Management, Aalborg University, Denmark

Introduction

Tele-healthcare has been growing rapidly because it promises fundamental benefits as improved access, reducing the cost of healthcare and increasing efficiency, improving quality and reducing travel time and related stresses for the patient [1].

This paper is about a multiple-case study in the development and implementation of two digitally supported health services in the treatment of patients suffering from alcohol addiction undergoing withdrawal symptoms and in patients who engage in sexual counselling.

Methods/Design

This study is a mixed methods multiple case study. Qualitative data include observations during the participatory design processes, semi-structured interviews with organizational health professionals and a review of documents (e.g., implementation and quality improvement plans, program manuals, etc.) that will shed light on the innovation process and specific implementation strategies that are used to implement the new intervention and practices. Additionally, focus groups with clinicians will explore their perceptions of the use of video-consultations [2,3,4].

We pursue a hypothesis that video consultations for and with patients suffering from alcohol withdrawal symptoms and people undergoing sexual counselling can empower the patient whilst freeing administratively bound resources to enable higher-quality care. The case-studies are based on theories of co-creation, specifically, we use elements from Participatory Design and User Innovation Management theory to co-create new, tele-medically inspired solutions that better patient treatment outcomes, compliance and engagement.

Results

The study shows that new technologies within healthcare can enhance public value particularly seen from a patient/user perspective. We also observe that the implementation of co-cre-ation derived methodologies has fostered a more innovative culture within the work-environments. The participatory design process with high user-involvement turned out to be a viable and feasible solution developing the video-consultations despite two clinical contexts marked by taboo and reluctance to deal with technologies.

Discussion

Technological innovation within healthcare does not always show cost effectiveness immediately. It is important to see development and implementation of tele-healthcare as a process that can improve the quality of care and eventually save money in the organisation. Working with tele-healthcare-innovation can, as in our cases, lead to changes within the culture in the organisation, so the innovation process create value in itself.

However, tele-healthcare in this field still poses some technical and practical problems for healthcare providers, such as worries about technology's limitations, lack of training and routinization. Yet, a faster and more responsive societal and political adoption curve and willingness to adopt video consultations on par with regular consultations is necessary, in order for this practice to become economically viable, sustainable and scalable for practitioners.

References:

[1] Fredskild TU, Dalkjær D, Langberg H. Innovation i sundhedsvæsnet. København: Gads Forlag, 2017

[2] Kanstrup A, Bygholm A, Bertelsen P, and Nøhr C. Participatory Design & Health Information Technology. IOS Press, 2017.

[3] Kanstrup A and Bertelsen P. User Innovation Management: A Handbook. Aarhus: Aarhus University Press, 2011.

[4] Clemmesen J, Rothmann MJ, Schmith AC, Caffery L, and Danbjörn DB. Participatory design methods in telemedicin research. Journal of Telemedicin and Telecare 2017: 23,9: 780-785.

Address for correspondence

Ulla Virkkunen Andrees, Ulla.Virkkunen.Andrees@rsyd.dk

Usability and Procedure Learnability of Evidence-based Interactive Clinical Systems: Roadmap for a Norwegian-Japanese Research Fellowship

Renée Schulz[a], Santiago Martinez[b], Takahiro Hara[c]

[a] *Department of Information and Communication Technology, University of Agder, Grimstad, Norway*
[b] *Department of Health and Nursing Science, University of Agder, Grimstad, Norway*
[c] *Department of Multimedia Engineering, Osaka University, Osaka, Japan*

Introduction and Background

Health information technology and information technology support for medicine are research priorities in Norway and Japan [1]. Japan is known for being one of the leading countries on innovative health technology developments, such as nursing robots. It was estimated that only the market for nursing robots would increase to ¥400 billion (around 30 billion NOK, 3 billion €) by the year 2035 [2]. This illustrates the innovation possibilities in the health technology sector. Norway has an above average rate of health service digitalization within the EU. Health technology was introduced into health education to further support professional health technology usage. In both countries, research and education in health technology is one of the pillars for society development and government priority.

Research and education on health information technology are closely connected in a way that latest developments should be introduced in education for next generation of health professionals, and theoretical principles (e.g., technology ethics, clinical and patient outcomes, safety) should be a core of health technology education to instruct the future generation of health technologists. Existing clinical procedures may be refined while new ones are introduced based on user needs and requirements. This situation requires health professionals and students to learn not only existing but new procedures as they are developed and used in clinical practice. However, the need of having to synchronise health education and practice regarding new procedures, technology and indicators demands an active learning where instructors (e.g., teachers, health professional in charge of a health department) and learners (e.g., health student, technology student, health intern, health professional) are supported by digital learning tools that allow them to quickly but effectively access procedure content and observe how procedures are operated in real life. To ensure that technology is accepted and usable for professional use, it must be integrated into the education process at the same time [3], which poses the challenge of learnability and continuous learning. In this scenario, the usability, accessibility and end-user contribution to digital health learning play a key role for the effectiveness, satisfaction and user experience of the learning process.

Methods

An interdisciplinary systematic research approach is proposed to cover aspects of information and communication technology, human-centred design, usability and user experience (UX) together with learning theories to support the learning and execution of clinical procedures. This includes the analysis and improvement of existing technology support within clinical workflows and education.

Table 1 – Research Questions and Targets

Research Question	Research Target
What are stakeholders' needs for digital technology support for health learning?	Stakeholders' needs
How can users contribute to the making and use of digital technology support for health learning?	End-user contribution
What methods are appropriate to analyse the effectiveness, efficiency and satisfaction of digital technology support for health learning?	User-centred design, Usability, technology adoption
How to handle the sources of clinical procedure and clinical outcomes to include them in digital technology support for health learning?	Big data

Expected Results

The aim is the evaluation, development and improvement of technology support used in specialized health services and health education. Additional results include the establishment of a cooperation between the Osaka University and the University of Agder to foster a participatory health technology education. Clinical outcomes and patient safety will be core of the assessment. These two factors are expected to improve with enhanced usability and UX of technology support.

References

[1] Veikart for bilateralt forskningssamarbeid med Japan, Forskningsrådet, 2014 [accessed: 21/05/2018]

[2] "Japanese Governmant to Urge Nursing Care Robot Development" (2016). ACCESSIBLEJAPAN.

[3] Nakamura, Y., Majima, Y., & Fukayama, K. (2016). The Usefulness of the Electric Textbooks at Nursing Practicam by Analyzing the Usage. *Nursing Informatics 2016*, 1038. [accessed: 21/05/2018]

Address for correspondence

Lead author: Dr. Renée P. Schulz, renee.schulz@uia.no