Nursing Informatics 2016
W. Sermeus et al. (Eds.)

doi:10.3233/978-1-61499-658-3-562

Integrating Social Media and Mobile Sensor Data for Clinical Decision Support: Concept and Requirements

Kerstin Denecke [1]
Bern University of Applied Sciences

Abstract. Social media are increasingly used by individuals for the purpose of collecting data and reporting on the personal health status, on health issues, symptoms and experiences with treatments. Beyond, fitness trackers are more used by individuals to monitor their fitness and health. The health data that is becoming available due to these developments could provide a valuable source for continuous health monitoring, prevention of unexpected health events and clinical decision making since it gives insights into behavior and life habits. However, an integration of the data is challenging. This paper aims triggering the discussion about this current topic. We present a concept for integrating social media data with mobile sensor data and clinical data using digital patient modelling. Further, we collect requirements and challenges for a possible realization of the concept. Challenges include the data volume, reliability and semantic interoperability.

Keywords. Social media, digital patient modelling, wearable sensors, decision support

1. Introduction

The individual state of a patient is very complex and concerns besides anatomy, physiology, metabolism, genetics also personal circumstances or life habits. Considering the individual state of the patient already in therapy planning is necessary for avoiding complications, predicting possible patient compliance, or for individualizing treatment decisions. Quantitative measurements as results from examinations are available in the electronic health record, described in clinical narratives or listed in terms of measured values. With an increase of information provision through social media and collected by patients individually through wearable sensors (e.g. fitness trackers), new sources of information reflecting patient's health observations and habits, i.e. quantitative and qualitative information are increasingly available. Activity trackers enable, low-cost wearable, non-invasive alternatives for continuous 24-h-monitoring of health, activity, mobility, life habits and mental status. The technologies have already been tested for their application in health monitoring systems, among others to analyse the tremor in Parkinson patients [1] or determine activity patterns for pulmonary patients [2] or analyse stress levels. Social media data was so far exploited for disease surveillance [9], patient recruitment or health

[1] Kerstin Denecke, University of Applied Sciences, Institute for Medical Informatics, Höheweg 80, 2502 Biel, Switzerland, kerstin.denecke@bfh.ch

communication. For example, studies showed that social media offers a medium to be used by the public, patients, and health professionals to communicate about health issues with the possibility of potentially improving health outcomes [8].

We envision the use of social media and wearable sensor data integrated with clinical data for continuous health monitoring, personalized treatment decisions and clinical decision support. As an example, consider the disease Multiple Sclerosis (MS), a disease that goes hand in hand with multiple symptoms (pain, mobility restrictions etc.) influencing massively the quality of life of patients. The entire care considers a mitigation of symptoms; healing is still unavailable. A continuous monitoring of MS patients could support in facilitating dealing with the situation by adapting the care plan accordingly targeting as hampering the quality of life as less as possible. However, we are still missing means that enable physicians to analyse and consider continuous health data as it could be collected from mobile sensors or from social media platforms within clinical decision making. In this paper, we introduce the concept of digital patient modelling as one possibility to include this non-clinical data in clinical decision-making. We present challenges and requirements towards an integration aiming at triggering the discussion of possible solutions.

2. Methods

2.1. Medical Social Media Data

With medical social media data we refer to web-based narrative text and data that contains medical content which was written by individuals (potential patients), physicians or other healthcare professionals. In *content communities and social networking sites* patients who suffer from diseases can share health data to empathise with each other or learn about experiences with treatments, physical exercises or medications. For example, PatientsLikeMe [https://www.patientslikeme.com/, last access: 17.11.2015] is a social network for patients that allows to share health-related experiences and compare treatments. Data is collected in this particular network partially in a structured manner: for the various features such as quality of life or single symptoms, categories are predefined (e.g. quality of physical life on a scale of 4 between best and worst). Similar to a paper-based diary, the authors of *weblogs or blogs* describe in unstructured format their personal opinions, impressions, feelings. In *online reviews*, individuals describe their experiences on symptoms, complications and effects of treatments with medical products including drugs.

2.2. Activity Tracker and Wearable Sensors

Smart wearable sensors are intelligent, low-cost, ultra-low-power sensor networks that enable individuals to collect huge amounts of biomedical information. There is a large variety of wearable sensors available including smart shirts, smart teeth etc.. We are concentrating in this paper on fitness or activity trackers as an example of wearable sensors (e.g. Microsoft Band, Fitbit Surge). However, challenges and requirements will be similar for other wearable devices. Activity tracker determine the activity progress in relation to a daily goal, distance and duration of activities, heart rate, calories burned, elevation, pace, sleep data and they even track food. APIs are provided to transfer the data from the fitness band into a database format, processable by computers (e.g. Apple

Health Kit, Google Kit, IFTTT - If This Then That). Challenges for processing data from fitness trackers are related to the volume of the collected data and the data quality. However, no comprehensive studies are already available on the accuracy of such sensor data.

2.3. Digital Patient Modelling for Exploiting Non-Clinical Data

In this section, we introduce digital patient modelling and the concept for integrating social media data and sensor data with clinical data in a digital patient model. We did not yet realized the concept since there are still several requirements and challenges to be addressed. They are summarized in section 3.

A digital patient model can be considered a specific and context-independent representation of a patient or of a specific disease or anatomical structure of a real patient, respectively. It consists of data elements that are semantically linked or grouped and thus provides an integrated view on the patient data reflecting his health status. Integrated in a decision support system digital patient models can support in therapy planning with individual quantitative optimization of clinical output or in predicting disease propagation. Treatment options can be simulated to support decision making. Two large EU initiatives worked towards a digital patient. Within Discipulus (http://www.digital-patient.net/), a roadmap, .i.e. a research agenda for the digital patient was formulated. The Virtual Physiological Human (VPH, http://www.vph-institute.org) is a framework of methods and technologies targeting at integrating fragmented health information. Once established, it will make possible the investigation of the human body as a whole. In this context, a sharing platform for biomedical data, tools, workflows and computing resources has been set up (VPH-Share). In terms of the model methodology, we can distinguish graphical patient models (based on radiologic image data) from probabilistic models. Clearly, social media and activity tracker data could be rather contribute to the latter, thus we are concentrating on this model type.

A frequently used method for knowledge modelling within clinical decision support systems are Bayesian networks which are probabilistic graphical models (e.g. Oniko et al. [4], Leibovici et al. [5]). Random variables represent in these graphs information entities such as medical examinations, medical imaging, patient behavior and patient characteristics (e.g. age, gender, tobacco and alcohol consumption), while directed edges represent the dependencies between information entities. Conditional probabilities need to be set for each information entity based on the graphical model structure. They define the correlation between an information and its direct causes. Bayesian networks allow for example to model dependencies among relevant information entities with respect to treatment decisions [6].

In previous work, we created a model for interdisciplinary treatment decisions related to laryngeal cancer [7]. The model represents a causal structure in the form of a Bayesian network. The graph is constructed using an overall structure of modelling diseases which was a result of the exemplary modelling. The underlying metastructure groups information items falling into the same group. For example, one subgraph integrates all information relevant to the tumor classification. Only clinical data is considered in the graph. Using this metastructure, additional subgraphs could be integrated to include data from social media and wearable sensors into this graph. This requires establishing a corresponding subgraph for reflecting in our case the relevant information from social media and activity trackers. More specifically, we suggest integrating one subgraph that comprises information on quality of life and behavior with data derived from medical social media. For example, this subgraph would contain nodes representing personal health perception, information on work, financial material wellbeing, personal safety, social relationships, emotional wellbeing, quality of environment, or hobbies (see Fig. 1). Another subgraph could reflect the fitness and activities. It integrates aggregated sensor data such as pulse, activity data, daily walking distance as a further extension. Nodes in theses graphs will be instantiated with data collected (filtered and analysed) from social media and the activity trackers.

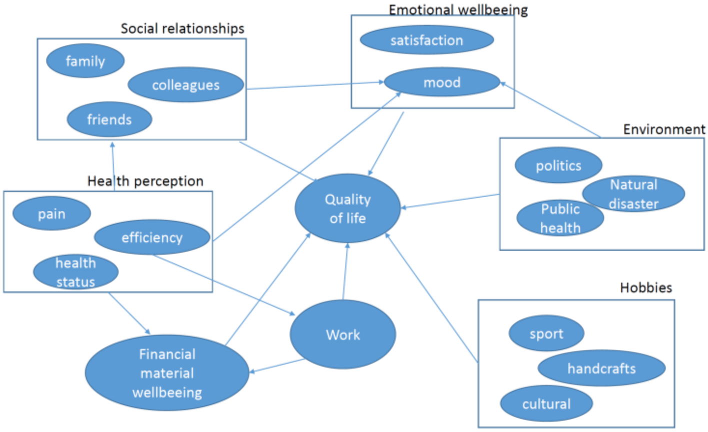

Fig. 1: Example subgraph for quality of life. Nodes of different categories represent the various factors that impact quality of life. For simplification, only few relationships are shown.

3. Requirements

Several requirements need to be addressed when realizing the above mentioned concept for clinical decision support. They can be grouped in three main categories.

1) Collection and analysis of data: Social media data and fitness tracker data need to be collected from the corresponding devices or sources. This requires actions from the patient side (e.g. posting information in a social media platform). Unstructured social media data needs to be processed, extracted and transformed into structured data that can be used to instantiate corresponding nodes in the patient model. Natural language processing methods are necessary to identify, analyse and process unstructured, free textual data, e.g. methods for extracting and categorizing relevant qualitative information as well as methods for linking diseases or symptoms to habits or treatments. Machine learning methods allow for automatic analysis, classification and clustering. Subjective information need to be interpreted, weighted and linked to objective clinical parameters. Fitness tracker and wearable sensors produce continuously data. To be useful, filtering and classification methods are required. This also includes methods for reasoning and inference to draw automatically conclusions and generate alerts. Methods need to be able to deal with data streams. Another challenge is semantic operability: it needs to be ensured that data from different sources can be integrated semantically in order not to lose or misinterpret information

2) Data quality: Exploiting social media and fitness tracker data in clinical decision making is only useful when data has a certain quality, i.e. is correct and time stamped. It should be possible to determine the identity of the person who is providing the social media and tracker information. Additionally, methods are required to check and weight the reliability of the data. Associated with the reliability of data is a critical interpretation and judgement of the data.

3) Modelling. Regarding the modelling, the graphical structure needs to be fixed for each subgraph, i.e. the structure of a quality of life subgraph and a fitness and activity graph needs to be set up. On the other hand, probabilities need to be assigned which is difficult given the fact that there are almost no experiences on correlations between different quality of life factors. In particular it is still unclear, how quality of life and fitness impact on treatment decision. Integration of the new subgraphs with the disease-specific subgraphs is another challenge.

4. Discussion and Conclusion

Through different social media sources and wearable devices, patient-collected clinical values (e.g. blood pressure, pulse, weight…), individual judgements on symptoms or efficacy of drugs and treatments, feelings and sentiments reflecting the health status are available. We suggest the concept of digital patient modelling as a mean to integrate the aggregated and filtered information from social media and sensors with clinical data. It is still unclear how this additional data impacts on decision making. An important issue is to be able to associate the data from social media and wearable sensors to a specific patient. The most efficient solution would be to use specific patient-doctor communication platforms for information provision. The data could then be directly collected from there and included into the model. In future work, the new subgraphs need to be developed and the usefulness of the additional data for patient care needs to be assessed. Once the patient-delivered data can be jointly analysed with clinical data, we are able to learn much more about influences of quality of life and fitness to the recovering process and to the patient's health.

References

[1] D.J. Wile, R.Ranawaya, Z.H.T. Kiss: Smart watch accelerometry for analysis and diagnosis of tremor. Journal of Neuroscience Methods 230 (2014): 1-4

[2] J. Juen, Q. Cheng, B. Schatz: A Natural Walking Monitor for Pulmonary Patients Using Mobile Phones. IEEE Journal of biomedical and health informatics, vol. 19, no. 4, 2015: 1399-1405

[3] K. Denecke: Health Web Science. Social Media Data for Healthcare. Springer International Publishing Switzerland, 2015

[4] Agnieszka Oniko and Marek J. Druzdzel. Impact of precision of bayesian network parameters on accuracy of medical diagnostic systems. *Artif. Intell. Med.*, 57(3):197–206, March 2013.L.

[5] L. Leibovici, M. Paul, AD. Nielsen, E. Tacconelli, and S. Andreassen. The treat project: decision support and prediction using causal probabilistic networks. *International Journal of Antimicrobial Agents*, 30 (Supplement 1):93–102, 2007.

[6] M. Cypko, M. Stoehr, K. Denecke, A. Dietz, and H.U. Lemke. User interaction with MEBNS for large patient-specific treatment decision models with an example for laryngeal cancer. *Int J CARS, Volume 9 (Suppl 1)*, 2014.

[7] M. Stöhr, M. Cypko, K. Denecke, HU. Lemke, A. Dietz. A model of the decision-making process: therapy of laryngeal cancer. Int J Comput Assist Radiol Surg. 2014;9(Suppl 1):217–8.

[8] Moorhead, S. Anne, et al. "A new dimension of health care: systematic review of the uses, benefits, and limitations of social media for health communication." *J of medical Internet research* 15.4 (2013).

[9] Velasco, Edward, et al. "Social Media and Internet-Based Data in Global Systems for Public Health Surveillance: A Systematic Review." *Milbank Quarterly* 92.1 (2014): 7-33.