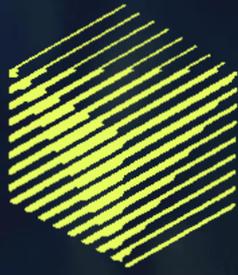


ERCIM



NEWS

www.ercim.eu

Special theme:

**Tackling Big Data
in the
Life Sciences**

Also in this issue:

Research and Society:

“Women in ICT Research and Education”

Can Data-driven Self-Management Reduce Low Back Pain?

by Kerstin Bach, Paul Jarle Mork and Agnar Aamodt

A new Horizon 2020 research and innovation project will start the development of the SELFBACK decision support system for self-management of low back pain in January 2016.

Low back pain is a common reason for activity limitation, sick leave, and disability. It is the fourth most common diagnosis (after upper respiratory infection, hypertension, and coughing) seen in primary care [1]. An expert group concluded that the most effective approach to manage non-specific low back pain is to discourage bed rest, to use over-the-counter pain killers in the acute stage if necessary (e.g., to be able to sleep), reassure the patient about the favourable prognosis, advise the patient to stay active, and advise strength and/or stretching exercise to prevent recurrence [2].

In January 2016 a new project funded by the European Commission's Horizon 2020 research and innovation programme will start the development of the SELFBACK decision support system for self-management of low back pain. The consortium includes the Norwegian University of Science and Technology (coordinator), the University of Glasgow and the Robert Gordon University in Aberdeen from the UK, the SMEs Kiolis from France and

Health Leads from the Netherlands as well as the National Research Centre for the Working Environment in Denmark and the University of Southern Denmark.

Self-management in the form of physical activity and strength/stretching exercises constitute the core component in the management of non-specific low back pain; however, adherence to self-management programs is poor because it is difficult to make lifestyle modifications with little or no additional support. Moreover, the way the self-management advice is presented, and whether the advice is followed up by reinforcement have substantial impact on adherence to a self-management plan and progress of symptoms. In the SELFBACK project we will develop and document an easy-to-use decision support system to facilitate, improve and reinforce self-management of non-specific low back pain. The decision support system will be conveyed to the patient via a smart-phone app in the form of advice for self-management. Even though there are about 300 pain-related

apps are available on the market, to the best of our knowledge, none of these apps have documented effects by scientific publications and none include an active decision support system. In contrast, we will document the effectiveness of the SELFBACK decision support system by conducting a nine-month randomized control trial.

The SELFBACK system will constitute a data-driven, predictive decision support system that uses the Case-Based Reasoning (CBR) methodology to capture and reuse patient cases in order to suggest the most suitable activity plans and goals for an individual patient. Based on personal preferences and clinical guidelines, the SELFBACK system will guide patients to adjust their lifestyle in order to optimize their management of low back pain, based on patient's activity data collected using a wristband. The incoming data will be analysed to classify the patient's activities and matched against the proposed activities in order to follow-up and advise the patient. The challenge of the data analysis is activity pattern detection, and matching it against existing patient profiles in order to suggest activity goals aiming at the most favourable outcome for the patient. Therefore we will enrich the data stream and the resulting patterns with the patient's symptom state, symptom progression and goal-setting.

Figure 1 gives an overview of the SELFBACK proposed data and decision process model that contains five processing modules, of which four modules target self-management by the patient while the fifth module targets the possibility for co-decision making between the patient and a clinician.

The initial screening module (1) acquires basic patient data, which is sent to the SELFBACK server. This data will be collected through a web-page and will provide the starting point for assessing the patient's status and for

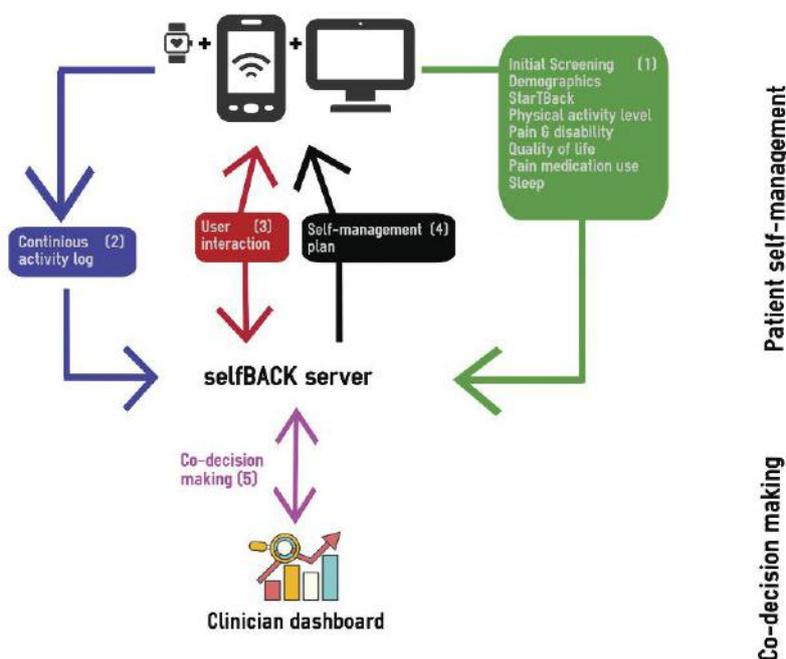


Figure 1: SELFBACK data and decision process model.

running the self-management planning module (4) for the first time. The plan for self-management will be updated and revised as more patient data is added. The activity logging module (2) runs continuously since it provides data from the wristband regarding the patient's physical activity and sleep patterns. The main goal of this module is to verify whether the patient follows the suggestions in the recommended plan (e.g., for physical activity). As for the initial screening data, the activity log data is also sent to the server, and becomes input to the periodic run of the self-management planning module (4). The user interaction module (3) is basically a question/answering module, typically initiated by the patient when using the SELFBACK app and browsing through the given information.

Overall, the aim of the SELFBACK system is to solve the problems of reinforcement and adherence to a self-management plan by offering an evidence-based system that allows personalized follow-up and advice to the patient.

Link:
SELFBACK project:
<http://research.idi.ntnu.no/selfback/>

References:
[1] P. Wändell et al.: "Most common diseases diagnosed in primary care in Stockholm", Sweden, in 2011. *Fam Pract* 30:506-513, 2013.
[2] M. van Tulder et al.: "Chapter 3. European guidelines for the management of acute nonspecific low back pain in primary care. *Eur Spine J* 15:s169-s191, 2006.

Please contact:
Kerstin Bach, Paul Jarle Mork, Agnar Aamodt
Norwegian University of Science and Technology
E-Mail: kerstin.bach@idi.ntnu.no,
paul.mork@ntnu.no,
agnar.aamodt@idi.ntnu.no

Twitter can Help to Find Adverse Drug Reactions

by Mark Cieliebak, Dominic Egger and Fatih Uzdilli

Drugs are great! We all need and use drugs every now and then. But they can have unwanted side-effects, referred to as "adverse drug reactions" (ADRs). Although drug manufacturers run extensive clinical trials to identify these ADRs, there are still over two million serious ADRs in the U.S. every year – and more than 100,000 patients in the U.S. die due to drug reactions, according to the U.S. Food and Drug Administration (FDA) [1]. For this reason, we are searching for innovative and effective ways to find ADRs.

Identifying ADRs is an important task for drug manufacturers, government agencies, and public health. One way to identify them before a drug goes to market is through clinical trials. Governments worldwide also have diverse surveillance programs in order to identify ADRs once the drugs are in use by consumers. For example, official websites such as MedWatch allow both patients and drug providers to submit ADRs manually. However, only a very small fraction of all ADRs is submitted to these systems – experts estimate that over 90% of all reactions go unreported.

Twitter can help!

On the other hand, there are millions of messages on Twitter that discuss medications and their side-effects. These messages contain data on drug usage in much larger test sets than any clinical trial will ever have. Inspired by this, research teams worldwide, including our team at Zurich University of Applied Sciences, are beginning to utilize these messages for ADR detection. The goal is

to automatically find relevant messages, to "understand" their content, and to extract structured data about the drugs and (unwanted) reactions.

A typical approach for ADR detection uses Natural Language Processing (NLP) to analyze tweets automatically. Input for the system is the entire stream of Twitter messages. Each individual tweet is analyzed, using a classification system as shown in Figure 1: The tweet is preprocessed and a set of relevant properties ("features") is extracted. Then, a classifier decides whether the tweet mentions an ADR. This classifier is based on machine learning and was trained beforehand on thousands of sample tweets that were tagged by humans. Finally, a system for named entity extraction is used to output a drug name and associated ADRs.

This approach is similar to technologies for sentiment analysis, which decide whether a tweet is positive or negative. Sentiment analysis is already success-

fully applied, for instance, in market monitoring, customer support and social media analysis.

State-of-the-art

Our ADR system, which implements the technologies shown above, achieves a success rate of 32% (measured in F1-score). This is comparable to other academic ADR systems: in an open international competition this year, even the best systems achieved only a success rate of approximately 40% [2].

For a preliminary evaluation on real-world data, we applied our ADR system to the full Twitter stream. The low precision of the system resulted in 20% of all tweets being classified as ADR. This is way too high; there are not that many ADR tweets in the Twitter stream. For this reason, we pre-filtered the stream with a list of 1678 drug names. Out of about 50 million tweets, this resulted in 13,000 tweets referencing drugs. Using the ADR system on this reduced set yielded 2800 tweets. We expect to find