KU LEUVEN

STADIUS
Center for Dynamical Systems, Signal Processing and Data Analytics

Citation/Reference	L. Billiet, T. Swinnen, R. Westhovens, K. de Vlam, S. Van Huffel, (2016), **SPARKLE@Home: Bringing Health Technology to Daily Life** International Conference on Engineering4Society, Leuven
Archived version	Author manuscript: the content is identical to the content of the published paper, but without the final typesetting by the publisher
Published version	NA
Journal homepage	http://www.engineering4society.org
Author contact	your email lieven.billiet@esat.kuleuven.be your phone number *+ 32 (0)16 327685*
IR	NA

(article begins on next page)

SPARKLE@Home
Bringing Health Technology to Daily Life

Lieven Billiet*†, Thijs Swinnen ‡§¶, Rene Westhovens‡§, Kurt de Vlam‡§, Sabine Van Huffel*†
*KU Leuven, Department of Electrical Engineering (ESAT-STADIUS)
†iMinds Medical Information Technology
‡University Hospitals Leuven, Division of Rheumatology
§KU Leuven, Department of Development and Regeneration,
Skeletal Biology and Engineering Research Center
¶KU Leuven, Department of Rehabilitation Sciences,
Musculoskeletal Rehabilitation Research Unit
lieven.billiet@esat.kuleuven.be

***Abstract*—In current clinical practice, physical therapy for rehabilitation purposes is still often practised solely in a clinical environment based on subjective or patient-reported assessment of activity capacity. This paper introduces SPARKLE, a project aiming at bringing physical therapy to the daily life. To this end, activity recognition and interpretable assessment based on mobile inertial sensors are being developed. The paper also discusses the challenges and opportunities in this interdisciplinary research project, as well as the importance of involving the patients in the development of the system.**

I. Introduction

Technology has an ever-increasing impact on many aspects of our lives, in areas as diverse as transportation, communication, production, education and many more. Among all these areas, health care is of particular interest. Here, technology not only opens new horizons (as in the aforementioned areas), but it aims at directly affecting a person's physical condition and well-being. Hence, the area should not focus on mere technical problems, but on the typical multidisciplinary amalgam that arises, involving engineers, clinicians and, most importantly, patients in the process.

Within healthcare, our research focuses on rehabilitation since it addresses a growing population, particularly among the elderly. Technology has a large impact in this area, mostly to provide clinicians with detailed information, but also in the way therapy is organized e.g. with exergames. For example, a lot of research involved so-called motion labs. Patients, e.g. athletes, are invited into a room equipped with stereoscopic cameras, force plates and other specialized equipment. Performing tests in such an environment allows the capture of full three-dimensional kinematic and kinetic objective data. It can be studied, summarized and compared among patients. Hence, it allows the clinician to not only rely on his or her experience for assessment and therapy planning, but also on objective knowledge obtained from data and the derived statistics [1].

Unfortunately, such an approach is expensive, cumbersome and artificial. A patient follows a predefined procedure in a clinical environment, whereas the movements in daily life are likely to be more important for the progress of the affliction. However, in the last decades, relatively cheap consumer electronics such as the Kinect, smartphones or other wearables allowed to move certain treatments from the hospital to the home environment. Examples include therapy sessions for elderly and patients [2], home monitoring such as fall detection [3] or studies of sedentary behaviour [4].

Although vision-based techniques remain present, particularly with the advent of Microsoft's Kinect, inertial sensors have gained popularity as well. They offer the advantage of not being limited to a field of view of a camera. As a result, they are more easily applicable for (long-term) monitoring and therapy integrated in daily life. Commercial examples of this include McRobert's MoveTest [5], for standardized activities, and Hocoma's Valedo [6], for exergames.

However, there is a downside to the use of technology in the medical world. Clinicians and therapists often encounter the information overload phenomenon: too much information cannot easily be merged into a meaningful diagnosis or prognosis. Clinical Decision Support systems have been developed to alleviate this issue. Driven by techniques of machine learning, they can be trained to recognize patterns from training data and apply the derived model on test data. Good results can be obtained, but often at the cost of transparency. Such 'black box models' are hard to interpret for the people involved. Particularly clinicians do not feel comfortable with blindly trusting the algorithm.

Our project, the development of a Sensor-based Platform for the Accurate and Remote monitoring of Kine(ma)tics Linked to E-health (SPARKLE) focuses on physical therapy for axial spondyloarthritis (axSpA) patients. This chronic disease is characterized by inflammation and ankylosis of the spine, limiting a patient's activity capacity (i.e. the ability to execute a task). Currently, activity capacity is assessed in the hospital by means of a subjective, patient-reported questionnaire and verified by therapists based on the execution of a set of informative transition activities. SPARKLE addresses several questions: how can we move this to a home environment? How can we obtain a more objective assessment and still

retain the important contribution of the patient? How can we communicate the related technical information in an interpretable way? How can this information be used to guide patients in their therapy? This paper discusses a part of SPARKLE's progress towards answers by focusing on a case study for activity recognition involving 28 axSpA patients. It also introduces an interpretable way to model activity assessment and discusses our approach to patient involvement.

The paper is structured as follows. The next Section explains the protocol and setup, followed by a Section that highlights the necessary activity recognition and one discussing activity assessment. The subsequent section focuses on the importance of communication towards clinicians, whereas the next stresses the importance of patient involvement and encouragement. Finally, the last Section discusses future work.

II. Protocol and Setup

The patient study was performed at the Division of Rheumatology of the University Hospital, Leuven [7]. The study protocol was approved by the Medical Ethics Committee (ML 5236).

Data collection involved 28 subjects (16 male, 12 female) with an average age of 43.7 years. The test protocol is inspired by the need to stay close to current clinical practice, taking into account the convenience for the patients. Therefore, the Bath Ankylosis Spondylitis Functional Index (BASFI) questionnaire was used [8], since it is already part of current clinical assessment. It defines a list of activities typically limited in axSpA patients. Each of them has an associated score, ranging from 0 to 10. In this way, a patient can indicate how apt he or she judges him/herself in performing the activity.

In our protocol, patients were first asked to fill out the questionnaire. The resulting BASFI scores range from 0/10 (best) to 8.1/10 (worst), with an average score of 3.14/10. It is a measure for the severity of activity limitations, albeit subjective. It reflects the self-judgement of the patient, often influenced by factors such as pain. As a next step, the patients were equipped with a two-axial accelerometer (Sensewear Pro 3 Armband, Bodymedia Inc, Pittsburgh, USA), sampled at 32Hz. This device was selected because it is easy to apply and comfortable to wear, in contrast to multisensor systems which require a more elaborate and precise setup. As a possible downside, only two-dimensional information of a single point of the human body is accessible. The Armband was mounted on the biceps of the dominant arm, its orientation matching the longitudinal and transversal axes. For a single-sensor system, the upper arm is a good location. On the one hand, it remains close to the trunk. This is useful to capture general body movements. On the other hand, it allows to capture movements of the arms as well, whilst not being subjected to larger and more variable movements as would be the case for i.e. the wrist. To further reduce the impact of arm movements, the patients were asked to fold the arms across the chest when possible. Subsequently, the patients performed a series of 10 activities inspired by the BASFI tasks. All of them are transitions, some of them repetitive, since these reveal most information on activity capacity. The activities are listed in Table I. The mounted sensor and some examples of a patient performing the activities are shown in Figure 1.

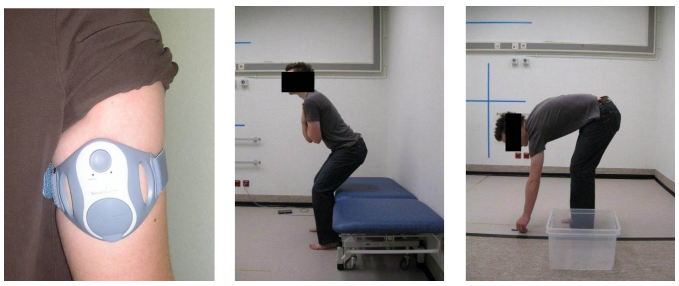


Figure 1. From left to right: the mounted sensor, STS activity and pen activity.

Table I
Description of the activities

Abbreviation	Description
getup	getting up starting from lying down
liedown	lying down starting from stance
maxreach	reaching up as far as possible
pen	picking up a pen from the ground
pen5	repeating pen 5 times, as fast as possible
reach5	touching a mark 5 times, as fast as possible
STS	performing a sit-to-stand movement
STS5	repeating STS 5 times, as fast as possible
sock	sitting down and putting on socks
stairs	climbing a stairs (13 steps)

Although the recordings took place in controlled conditions in the hospital, the protocol allows for portable data collection at home. In fact, more recent data acquisitions already took place in the home environment. It remains to be seen whether this influences the results.

III. Activity Recognition

The output of the data acquisition are continuous acceleration signals. In order to gather useful information from the continuous recordings, activities first need to be segmented and recognized.

A. Segmentation

Due to the setup of the experience, a simple segmentation approach is possible. The patients remained static in between activities. Hence, bursts of higher variability indicate the regions of interest. A semi-automatic thresholding algorithm was previously developed by M. Milosevic [7]. It is important to notice it would not be sufficient for more general segmentation of activities in daily life situations. Currently, other approaches are being investigated to replace this step. Yet, in this controlled setup, the algorithm succeeds in separating the segments containing the activities of interest from the non-informative static regions.

B. Recognition

Throughout activity recognition literature, two main approaches often appear.

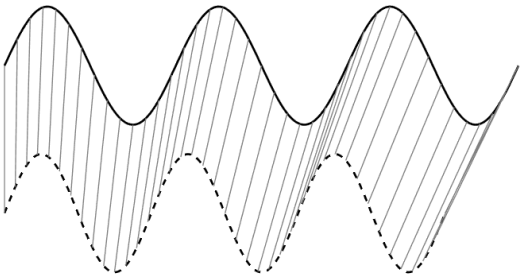


Figure 2. Example of DTW matching of two sinusoids with slightly different frequencies.

a) Sliding windows: A first approach uses windows of fixed or variable length. Without the need for segmentation, the signal is divided in (sometimes overlapping) parts. Each of these parts can be characterized by features. Some are simple statistical measures such as the mean or variance, others are more complicated, but still in the time domain. Another possibility is feature extraction from the frequency or wavelet representation [9]. The approach has been applied with high accuracy, but it has some disadvantages. It is mostly suitable for static or repetitive activities such as sitting, standing or walking. Consequently, transitory movements are more difficult to detect. Moreover, fixed windows do not allow a precise segmentation. Hence, the technique is often applied for long term monitoring, where crude information suffices, e.g. in studies of the effects of a sedentary life style [4].

b) Pattern matching: Features extracted from sliding windows can serve as inputs for a classifier, but they are not always easily interpretable. On the other hand, a second approach to the activity recognition problem directly relates to the acceleration signals. In pattern matching, one first constructs a template for each activity based on training examples. Then, potential activities can be matched against the template. The advantage of this method is that it corresponds to how human observers such as clinicians deal with the data: shapes identify activities. Hence, the method is interpretable. Furthermore, patterns are better suited to capture transitions. Yet, it is hard to build a reliable pattern that captures the variability in the examples, an evident drawback with respect to human observers. This can be alleviated to some extent by a technique known as dynamic time warping (DTW), illustrated in Figure 2. Pattern matching has been applied for recognition of sit-to-stand [10].

c) A combined approach: Recently, we performed a study showing that, for segmented data, the two approaches are complementary. Combining them outperforms the individual approaches with statistical significance, leading to a recognition rate of 93.6% in a leave-one-subject-out validation [11]. As window features, we selected among others common signal processing measures that reflected the signal level (e.g. mean), its energy (power) and its self-similarity (autocorrelation). Each of the chosen features can be linked to properties of the signal. Additionally, the similarities to the ten activity templates were added to the feature vector. This practice is known as early fusion. Finally, classification was performed with a linear classifier. Figure 3 shows a graphical representation of the final algorithm.

IV. Activity assessment

The next part of the workflow is the assessment of the recognized activities. Although this has not yet been applied to the current study, we already developed a framework for interpretable feedback based on performance indicators extracted from the data. As mentioned in the introduction machine learning readily allows to extract information from data and to provide predictions based on underlying data patterns. Yet, standard techniques such as Support Vector Machines are mostly black boxes, hence difficult to interpret. Another kind of decision support has been used frequently in medical practice: scoring systems. They list a number of criteria for measured values, associating a cost to all of them. In the end, a total score can be calculated as a simple sum of all relevant costs. This score can be mapped to an empirical risk. Examples of such scoring systems include CHA_2DS_2-VASc for predicting the risk on atrial fibrillation or ASDAS to assess disease activity in Ankylosing Spondylitis, a subgroup of axSpA exhibiting definite radiographic bone formation at the sacroiliac joints [12]. Such scoring systems are powerful rules of thumb, sometimes tied to concrete treatment guidelines. They summarize the experience and knowledge of clinicians. Yet, they are often rather ad-hoc, though sometimes validated by subsequent statistical analysis.

Our approach tries to combine the power of machine learning with the interpretability of scoring systems. More concretely, our Interval Coded Scoring (ICS) constructs a scoring system from training data [13] using techniques from machine learning and optimization theory. The scoring system results from solving the following problem (in matrix notation):

$$\min_{w,b,\epsilon} \|Dw\|_1 + \gamma\epsilon^T\mathbf{1} \tag{1}$$
$$\text{s.t.:} \begin{cases} Y(Zw+b) \geq \mathbf{1} - \epsilon \\ \epsilon \geq \mathbf{0} \end{cases}$$

The first constraint is related to correct classification, in which the ϵ values are slack variables with regard to the classification margin. Hence, the objective function balances a sparsity objective ($\|Dw\|_1$), with the permissible error on the classification of the training data ($\epsilon^T\mathbf{1}$) using the hyperparameter γ.

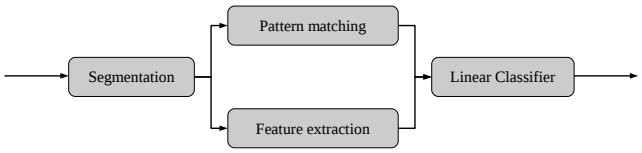


Figure 3. Overview of the activity recognition algorithm

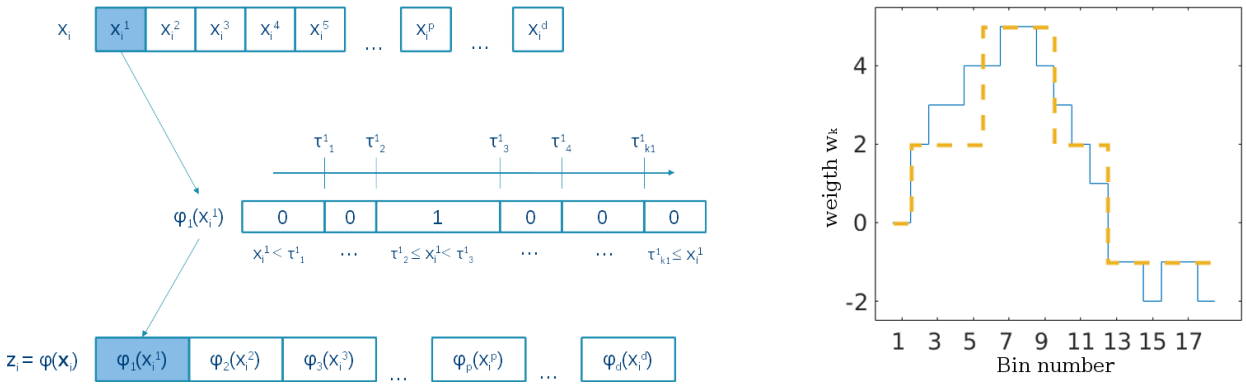


Figure 4. ICS data transformation (left) and the influence of sparsity (right). Non-sparse optimization is in full, the sparse result is shown with a dashed line.

Two important aspects are illustrated in Figure 4. The training data Z is binary. It is obtained by binning the original data X. Each of the bins gets a corresponding weight gathered in the weight vector w. This is shown in the left part of the Figure. The right part shows the impact of $\|Dw\|_1$. D is a difference matrix, describing the differences between weights of adjacent bins. Requiring these differences to be sparse creates a more simple scoring system, less susceptible to noise.

The weights can be converted to costs associated with each bin. Summation of these costs (equivalent to calculating Zw) yields the final score. As in an ordinary scoring system, the score can be mapped to a risk by a logistic regression. An example of the application of the approach for the vertebral column data set (UCI database) is shown in Figure 5. The first three lines show the selected indicators and the associated costs of their intervals (bins). The last line shows the risk of vertebral disease for each possible score.

Currently, the method has not yet been applied for activity assessment. It would first require the selection of potential features of interest based on a thorough discussing with physical therapists. Possible features could include the duration of (parts of) an activity, the slope of the acceleration etc. Moreover, in the current version, the score values are theoretically unbounded. These values should be mapped to a standardized quality assessment measure, rather than a risk. Such a standardized measure could be akin to the BASFI score currently used, but using the patient's subjective self-assessment as one of the inputs rather than the only input.

The representation can easily be used by therapists and patients alike, explicitly indicating why an execution of a certain activity corresponds to a good or bad activity capacity. When tracked over time, it also shows which of the contributing factors improves or deteriorates, with or without consequences. Because the classifier is transparent, a patient is more involved and can express his/her own opinion more confidently.
The method is objective, hence satisfying the needs of therapists, but it is trained using both objective measure and the patients' subjective assessment. Needless to say that it is not meant as a replacement for the current system, but as an additional source of information: patients should still be allowed to convey their experiences to the therapists during check-up sessions.

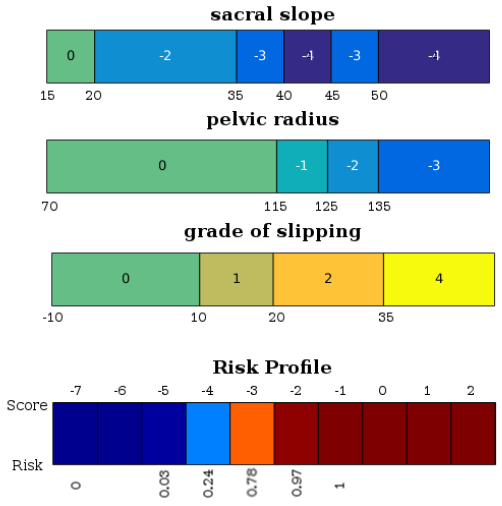


Figure 5. An example of an ICS model (UCI vertebral column dataset)

V. Communication and Collaboration

SPARKLE is a multidisciplinary endeavour. This poses several challenges since different fields tend to use different techniques, nomenclature and approaches. As an example, data collected in the University Hospital tends to be exported in excel files. Statistical analysis is performed with specialized software packages such as SPSS. On the other hand, the signal processing research group develops its algorithms in Matlab.

One model of collaboration is a strict separation. In this case, one could argue that the Hospital takes care of patient studies and delivers the data. Data analysis happens in the signal processing group and results are transferred back to the Hospital for assessment. Indeed, several smaller studies have been performed in this way. However, this leads to a large communication overhead in case of problems. Moreover, captured signals need to be segmented manually to obtain training data for recognition. Therapists have the expertise to do so, but often lack the software.

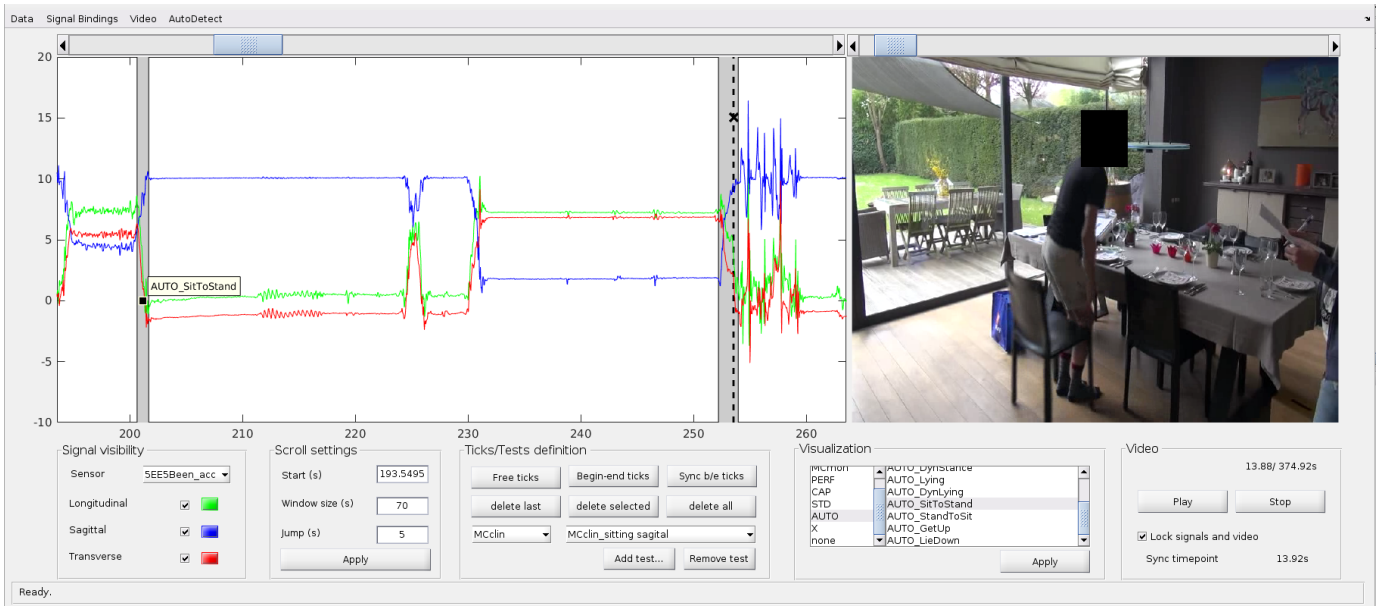


Figure 6. Graphical interface for segmentation, annotation and recognition. The highlighted zones are automatically recognized sit-to-stand transitions.

Therefore, we developed a graphical interface. For the time being, it requires Matlab to be installed. In the future, an implementation in a non-commercial language will be developed. Whatever the language, the interface creates a layer of abstraction accessible and interpretable to all backgrounds involved. It unifies import, visualisation, segmentation of data and automatic recognition of activities. It was developed taking into account feedback and requirements of both parties involved. The interface is presented in Figure 6. At the time of writing, it offers data import from several formats, manual segmentation of activities, the possibility of video synchronized signal visualization, sensor and axis selection, visualization of previously segmented activities, automatic recognition of pose and pose transitions, saving all changes and several smaller support options.

VI. Involving the patient

The system described in this paper can be used in several ways. On the one hand, it can serve as a source of information for therapists. In that sense, it is an assessment and follow-up tool helping the therapist to provide the patient with adequate therapy. On the other hand, SPARKLE also focuses on the patients themselves, in collaboration with the e-Media Lab (Faculty of of Engineering Technology, Campus Group T, Leuven). Based on the directives of the therapist, patients are expected to perform certain exercises at home. Adherence to this therapy is a known problem, particularly if exercises lead to a certain amount of discomfort. Furthermore, even though patients may be willing to perform the exercises, they are unsupervised. For the therapist and patient alike it is hard to know whether the exercises were performed correctly. Hence, these two issues, adherence and feedback, need to be addressed.

These issues can be dealt with by an appropriate design of the system presented to both patient and therapist or clinician. Because it is meant to be used in a home environment, a mobile application seems the most appropriate. To encourage the patient to use the system, the principles of *Persuasive Design* [14] should be applied: how does one convince the user of the usefulness of the application and/or encourage him or her to continue using it? These principles strongly overlap with those attributed to *gamification*. Gamification is defined as 'the use of game design elements in a non-gaming context' [15]. It points to both the elements, e.g. badges, experience points, score boards, ... as well as the game dynamics that are essential to a game experience, e.g., providing clear goals, offering a challenge, showing progress, providing feedback, giving a story or theme etc [16]. Hence, gamification is about harnessing the motivational affordances of gameful experiences to influence psychological outcomes and further behavioral outcomes [17]. In this case, it would motivate patients to greater therapy adherence and offer them fast and effective feedback on their activity capacity.

Of course, as for any system, this only works if the design is tailored to the needs and concepts of all stakeholders: patients, clinicians and therapists. Therefore, it is important to include them in the design process from the start. Such an approach is called *Participatory Design* (PD). PD is 'a set of theories, practices and studies related to end-users as full participants in activities leading to software and hardware computer products and computer-based activities' [18]. In PD, each stakeholder is recognized as an authority in his or her own domain. All stakeholders should be considered co-designers in all stages of the design process. This allows to discuss the aspects of the design process on (near) equal footing in an attempt to eliminate a bias to either the technical, medical or usability

side: all of them need to be considered to end up with a successful system [19]. In case of SPARKLE, this is enforced by following the P-III methodology and process [20]. The process consists of several stages: contextual inquiries to get to know the patients' home environment, focus groups and co-creation sessions, formative assessment and finally adoption studies. Currently, patient groups and health professionals are being recruited for the focus groups.

VII. Conclusions and Future Work

In this paper, we presented SPARKLE, a multidisciplinary project on recognition and assessment of activities as support in physical therapy. To this end we focused on a case study. Although the case study was performed in a clinical environment, it stressed the importance of using wearables and choosing a set of activities that can easily be performed in an unsupervised way in the home environment. We presented a successful approach to activity recognition, followed by a description of an interpretable scoring system that can be used for assessment. Then, we elaborated somewhat on the problems related to communication and collaboration across different fields. Finally, we discussed the importance of involving the patient throughout the project. We proposed the concepts of Persuasive Design and Participatory Design to allow integration of the technical aspects in a mobile application offering adequate feedback whilst encouraging the patient to adhere to the therapy. Throughout this paper, some future work has already been outlined.

A first step has already been taken. Data is now being collected in the home environment. Although the protocol remains the same, it is no longer performed in the perceived controlling environment, effectively bringing the technology to the patients' homes.

Secondly, the assessment step should be applied to the specific problem. This involves specification of clinically relevant indicators to be used for ICS, but also specifying the target values in collaboration with clinicians and patients. This last merging step is essential to combine subjective assessment with objective measures and clinical experience.

Finally, once the full workflow at home and the mobile application have been developed and tried, a larger-scale study should assess the system's validity.

VIII. Acknowledgements

Bijzonder Onderzoeksfonds KU Leuven (BOF): Center of Excellence (CoE) #: PFV/10/002 (OPTEC), SPARKLE – Sensor-based Platform for the Accurate and Remote monitoring of Kinematics Linked to E-health #: IDO-13-0358; iMinds Medical Information Technologies: Dotatie-Strategisch basisonderzoek (SBO- 2016); Belgian Federal Science Policy Office: IUAP #P7/19/ (DYSCO, 'Dynamical systems, control and optimization', 2012-2017); EU: European Union's Seventh Framework Programme (FP7/2007-2013): ERASMUS EQR: Community service engineer , #539642-LLP-1-2013; EU: The research leading to these results has received funding from the European Research Council under the European Union's Seventh Framework Programme (FP7/2007-2013)/ERC Advanced Grant: BIOTENSORS (n° 339804). This paper reflects only the authors' views and the Union is not liable for any use that may be made of the contained information.

We would like to thank prof. L. Geurts and prof. V. Vanden Abeele (e-Media Lab, Faculty of Engineering Technology, Campus Group T, KU Leuven) for their contribution and helpful feedback.

References

[1] H. Zhou and H. Hu, "Human motion tracking for rehabilitation—a survey," *Biomedical Signal Processing and Control*, vol. 3, no. 1, pp. 1 – 18, 2008. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S1746809407000778

[2] A. Kramer, C. Dettmers, and M. Gruber, "Exergaming with additional postural demands improves balance and gait in patients with multiple sclerosis as much as conventional balance training and leads to high adherence to home-based balance training," *Archives of Physical Medicine and Rehabilitation*, vol. 95, no. 10, pp. 1803 – 1809, 2014. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0003999314003402

[3] S. Abbate, M. Avvenuti, F. Bonatesta, G. Cola, P. Corsini, and A. Vecchio, "A smartphone-based fall detection system," *Pervasive and Mobile Computing*, vol. 8, no. 6, pp. 883 – 899, 2012, special Issue on Pervasive Healthcare. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S1574119212000983

[4] P. A. Semanik, J. Lee, J. Song, R. W. Chang, M.-W. Sohn, L. Ehrlich-Jones, B. E. Ainsworth, M. M. Nevitt, C. K. Kwoh, and D. D. Dunlop, "Accelerometer-monitored sedentary behavior and observed physical function loss." *American Journal of Public Health*, vol. 105, no. 3, pp. 560–566., March 2015.

[5] McRoberts, "MoveTest: Unobtrusive assessment of Physical Function under supervised conditions." http://www.mcroberts.nl/products/mt/, accessed April 6, 2016.

[6] Hocoma AG, "Back pain therapy with the Valedo THERAPY Concept," https://www.hocoma.com/world/en/products/valedo-therapy-concept/, accessed April 6, 2016.

[7] T. Swinnen, M. Milosevic, S. Van Huffel, W. Dankaerts, R. Westhovens, and K. De Vlam, "Proof of concept : the novel instrumented Bath Ankylosing Spondylitis Functional Index (iBASFI) shows promising reliability and validity in the assessment of activity limitations in axial spondyloarthritis," *Journal of Rheumatology (accepted)*, 2016.

[8] A. Calin, S. Garrett, H. Whitelock, L. G. Kennedy, J. O'Hea, P. Mallorie, and T. Jenkinson, "A new approach to defining functional ability in ankylosing spondylitis: the development of the bath ankylosing spondylitis functional index." *Journal of Rheumatology*, vol. 21, no. 12, pp. 2281–2285, 1994.

[9] S. Preece, J. Goulermas, L. Kenney, D. Howard, K. Meijer, and R. Crompton, "Activity identification using body-mounted sensors – a review of classification techniques," *Physiological Measurement*, vol. 30, pp. R1–R33, 2009.

[10] R. Ganea, A. Paraschiv-Ionescu, and K. Aminian, "Detection and classification of postural transitions in real-world conditions," *Neural Systems and Rehabilitation Engineering, IEEE Transactions on*, vol. 20, no. 5, pp. 688–696, Sept 2012.

[11] L. Billiet, T. Swinnen, R. Westhovens, K. de Vlam, and S. Van Huffel, "Activity recognition for physical therapy - fusing signal processing features and movement patterns," in *3rd International Workshop on Sensor-based Activity Recognition and Interaction (iWOAR2016)*, 2016.

[12] C. Lukas, R. Landewé, J. Sieper, M. Dougados, J. Davis, J. Braun, S. van der Linden, and D. van der Heijde, "Development of an asas-endorsed disease activity score (asdas) in patients with ankylosing spondylitis," *Annals of the Rheumatic Diseases*, vol. 68, pp. 18–24, 2009.

[13] L. Billiet, S. Van Huffel, and V. Van Belle, "Interval coded scoring index with interaction effects: a sensitivity study," in *5th International Conference on Pattern Recognition Applications and Methods (ICPRAM2016)*, 2016, pp. 33–40.

[14] J. Geuens, L. Geurts, T. Swinnen, R. Westhovens, K. De Vlam, and V. Vanden Abeele, "A Systematic Review of the Persuasive Design Principles Used in Health Apps for Arthritis," in *Annual Meeting of the American College Of Rheumatology (ACR/ARHP)*, November 2015.

[15] S. Deterding, D. Dixon, R. Khaled, and L. Nacke, "From game design elements to gamefulness: Defining "gamification"," in *Proceedings of the 15th International Academic MindTrek Conference: Envisioning Future Media Environments*, ser. MindTrek '11. New York, NY, USA: ACM, 2011, pp. 9–15.

[16] B. Cugelman, "Gamification: What it is and why it matters to digital health behavior change developers," *JMIR Serious Games*, vol. 1, no. 1, p. e3, Dec 2013. [Online]. Available: http://games.jmir.org/2013/1/e3/

[17] J. Hamari, J. Koivisto, and H. Sarsa, "Does gamification work? – a literature review of empirical studies on gamification," *2014 47th Hawaii International Conference on System Sciences*, vol. 0, pp. 3025–3034, 2014.

[18] M. J. Muller, "Participatory design: The third space in hci," in *The Human-computer Interaction Handbook*, J. A. Jacko and A. Sears, Eds. Hillsdale, NJ, USA: L. Erlbaum Associates Inc., 2003, pp. 1051–1068. [Online]. Available: http://dl.acm.org/citation.cfm?id=772072.772138

[19] T. Bratteteig and I. Wagner, *Disentangling Participation: Power and Decision-making in Participatory Design*. Springer Publishing Company, Incorporated, 2014.

[20] V. Vanden Abeele, B. De Schutter, L. Geurts, S. Desmet, J. Wauters, J. Husson, L. Van den Audenaeren, F. Van Broeckhoven, J.-H. Annema, and D. Geerts, *P-III: A Player-Centered, Iterative, Interdisciplinary and Integrated Framework for Serious Game Design and Development*. Berlin, Heidelberg: Springer Berlin Heidelberg, 2012, pp. 82–86.