Ecologically valid trials of elderly unobtrusive monitoring: analysis and first results

Antonis S. Billis[1], Panagiotis Kartsidis[1], Dimitris-Konstantinos G. Garyfallos[1], Marianna S. Tsatali[1], Maria Karagianni[1] and Panagiotis D. Bamidis[1]

[1]Medical Physics Laboratory, Medical School, Faculty of Health Sciences, Aristotle University of Thessaloniki
{ampillis, bamidis}@med.auth.gr
{panos.kartsidis, mkaragianni.psy}@gmail.com
dgaryfal@physics.auth.gr
mtsatali@yahoo.gr

Abstract. Intelligent health monitoring systems of elderly have been around for several years now. Evaluation of sensor measurements and intelligent processing algorithms has been performed mainly in lab settings, prohibiting the collection of datasets that reflect real behavior of seniors. As a result, when technology migrates to real-life settings, fails to achieve similar monitoring accuracy. Our approach tackles this problem, by piloting the USEFIL intelligent monitoring system, to elderly people both at lab and home settings. Fifteen (15) seniors were recruited to follow a number of predefined activities in a free-form manner for 2 weeks. Five (5) of them were also recruited for piloting the system in their own homes for a period of two months. Statistical analysis of sensor observations and clinical assessment tools revealed the monitoring added value of the sensors in an ecological valid environment. In addition, trend analysis based on lab findings, showed – by means of a single case study- the potential of the system to continuously assess health indicators and detect health deterioration signs.

Keywords: ecological validity; continuous in-home health assessment; active and healthy ageing; statistical process control; living lab; ambient assisted living

1 Introduction

Ambient Assisted Living (AAL) systems have widely developed and evaluated towards their capacity to monitor pathological patterns in elderly people, so to promote early risk identification, related to chronic diseases [1], [2]. However, most approaches followed have severe limitations in their prospect to be applied in real-life settings [3], since evaluation of algorithms is done either by recruiting young adults [4] or by strict lab experiments [3] or short-term trials at home with small amount of trial homes [5]. Our approach provides evaluation of the USEFIL intelligent monitoring system [6], both in an ecologically valid lab environment and at seniors' residencies. Analysis of low level events, derived by sensors, are correlated to clinical assessment batteries, providing evidence for the clinical added value of the USEFIL intelligent monitoring system. Contrary to existing work in the field [3], free-form activities have been introduced to alleviate strict execution of tasks, resulting to a free-form, ecological valid dataset. Long-term, trend analysis has been subsequently applied to low-level events that have been found statistical significantly correlated to clinical assessment tests. Statistical process control modeling [7] has allowed for retrospective visualization of seniors health patterns, while leaving at their own homes.

2 Materials & Methods

2.1 Lab pilots

Lab pilots ran in Thessaloniki, in the Active & Healthy Aging Living Lab (AHA LL). There, a living room environment and a kitchen environment were set up in the same room. The initial layout of the AHA LL is visualized in Fig. 1. In order to look more realistic, AHA LL was equipped with home appliances and furniture so as to better resemble a senior's home. There, the necessary technological infrastructure and the USEFIL hardware were installed.

Fig. 1. AHA LL spaces & monitoring system unobtrusive set up

The methodology that was followed towards the execution and evaluation of the trials at the lab was: i) recruitment, ii) baseline assessment & follow up, iii) protocol of directed activities definition, iv) trial execution – ongoing period of trials, v) end users feedback and vi) data analysis.

As a first step seniors' demographic data and medical history were obtained. Global cognitive functioning was assessed using the MMSE. Depression levels were evaluated with the PHQ-9 scale, Quality of life index was measured by SF12, ICECAP and ASCOT INT 4, whereas the ability of independent living was assessed by the Barthel index. Fullerton test was used to assess participant's physical performance. After a two weeks period, participants were assessed to the previous assessment battery for follow up purposes.

The real testing and use of the environment took place for 8 days maximum for each participant. Each session lasted approximately 60-90 minutes. The participants were asked to conduct a series of specific tasks as independently as possible.

Fig. 2. Recorded activities in AHA LL

2.2 Home pilots

Technical setup of the USEFIL system took place in five (5) seniors' homes. USEFIL software and hardware was installed and setup a-priori at lab premises. Typical installation example is shown in **Fig. 3**.

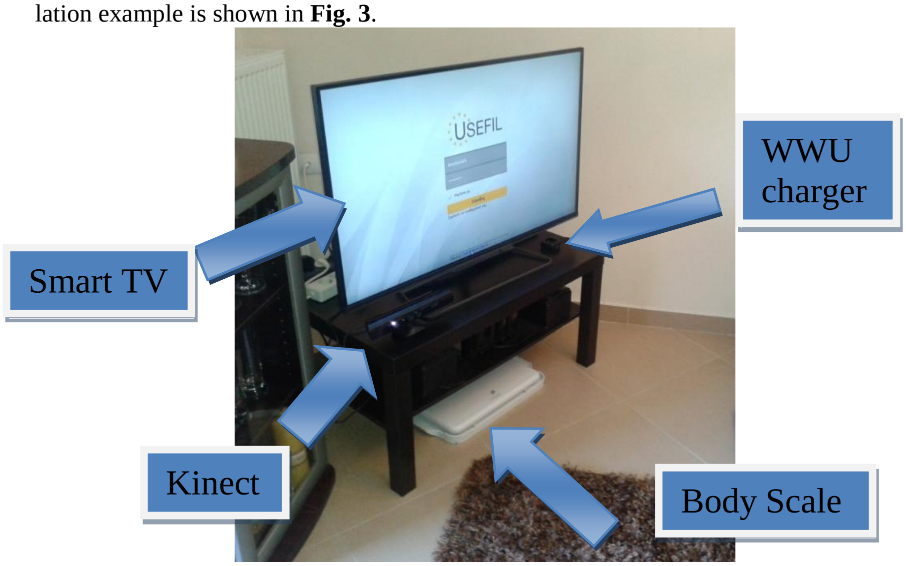

Fig. 3. USEFIL system setup at senior's home

Five (5) elderly, lone-living women aged 75,6±4,72 years and 14.8±6.57 years of education were recruited. Four out of five seniors (4/5) had memory problems, while

two (2) of them had depressive symptomatology. All five seniors had participated in the lab pilots. So, recruitment took place after they had completed the testing of the system and were asked of their intention to use the USEFIL system at their own homes, in the realm of a focus group discussion. Participants, that declared interest, were explained about the purposes of the home study and upon acceptance, they signed an informed consent, declaring their voluntary participation. Seniors were examined by two (2) neuropsychologists at baseline, one-month follow up and at the end of the two-month period. Global cognitive functioning was assessed using the MMSE. Depression levels were evaluated with the PHQ-9 scale, Quality of life index was measured by SF12, ICECAP and ASCOT INT 4, whereas the ability of independent living was assessed by the Barthel index. Fullerton test was used to assess participant's physical performance.
After, the initial training period neuropsychologists either visited in person seniors twice per week or they contacted them via telephone. Seniors were encouraged to perform a list of minimum optional daily tasks related to their interaction with USEFIL system's devices and apps.

3 Results

3.1 Sensors vs Clinical assessment

In order to evaluate the clinical added value of the USEFIL system, sensor measurements (Low-Level Events) were correlated to the battery tests that were performed at the baseline and the follow up. In particular, correlation analyses were performed between the neuropsychological, physical test results and sensors' observations. The correlation coefficient used was Pearson's r. The statistical significant findings of the analyses are shown in Fig. 4. PHQ results – which refer to the assessment (existence and severity) of the depressive symptomatology - were correlated either negative or positive to mobility or gait parameters as measured by sensors, e.g. StepCount (number of Steps per minute), WalkingSpeed (cm/sec), feetElevation (height of feet while walking in front of the Kinect) speech and facial expression characteristics, e.g. speech arousal, eyes' blinking rate and facial skin color redness level. Most of the above findings are in line with medical literature [8][9][10]. SF12 mental component is a subjective feeling of a senior about his/her mental ability/ies. This subjective measure of quality of life was negatively correlated to feet elevation. However no data are available, supporting the fact that someone has increased levels of quality of life, while their feet elevation decreases. A statistically significant relationship was found between ICECAP (sum score) and walking speed ($p=.046$). This evidence is in line with previous studies [11], where walking speed is considered as a predictor of quality of life. Additionally, the item of 'thinking about the future' from ICECAP is related to speech arousal ($p=.05$), which means that participants who expressed worries and were anxious regarding the future, were more likely to have higher speech arousal scores, compared to those who felt more safe about the future. The variable of independence, measured by ICECAP too, is related to the sitting speed and the walking speed, which means that those who feel independent in their daily life, had a better mobility status. Furthermore, ASCOT INT 4 scale, which also assess quality of life, was found to be positively related to speech arousal ($p=.015$), and negatively related to sitting speed (p=.034). Although there are no data supporting this evidence, it is a quite important evidence to be studied in following studies. Parameters of independent living, specifically, bowels control, toilet use and transfer activity, are significantly related to sitting speed, while toilet use (p=.025) and transfer activity (p=.025) are correlated with number of steps. Chair stand test (measures lower body strength in terms of number of completed chair stands in 30 seconds) was negatively related to feet elevation, step count and sitting speed. Lower sitting speed time denotes better balance and lower body strength. Therefore more repetitions executed by participants show their good balance ability and lower body strength. 2-minute step test (measures seniors' aerobic endurance and dynamic balance) was negatively related to sitting speed and walking speed. The latter seems to be inconsistent and it needs more data to be confirmed. Finally, Foot up & go test (measures speed, agility and balance while moving) had statistical significant relationship to step count (p=.044), while it is negatively related to walking speed (p=.036). The latter was an unexpected result and needs to be studied with more participants.

	FeetElevation	SittingSpeed	StepCount	WalkingSpeed	Emotion	SpeechArousal	FaceAngry	FaceBlinkingRate	FaceHR	FaceRed
PHQ- 7 (concentration deficits)			-.635* .036	-.716* .020		-.515* .049				
PHQ-8 (agitation/retardation)			-.771** .005	-.697* .025						
PHQ-9 (2)	.833* .020							.592* .042		-.641* .025
PHQ-2 (depressive mood)	.813** .004									
PHQ-4 (loss of energy)	.692* .026							.669* .017		-.568* 0.0271266
MMSE					-.759* .011				.790** 0.0007787	
Chair Stand Test	-.919** .003		-.711* .048	-.774* .041						
2-Minute Step Test		-.645* .032		-.823* .012						
Foot Up & Go Test 2			.681* .044	-.741* .036						
SF (MCS)	-.819** .004									
ICECAP-O (sum)				.640* .046						
BOWELS		-.616* .025								
TOILET USE		.a 0.000		.697* .025						
TRANSFER		.a 0.000		.697* .025						
INT4. Current SCRQol		-.590* .034				.616* .015		-.666* .018		
Independence		.a 0.000		.848** .002						

Fig. 4. Clinical Assessment vs Sensor Measurements correlation (1st line – Pearson correlation, 2nd line - significance)

3.2 Long-term follow up

In order to demonstrate the monitoring capabilities of the USEFIL system within home settings, long-term trends relative to health parameters such as mobility, gait, emotion and cognition. Taking into account lab findings, i.e. correlations found among sensor observations and clinical assessment batteries, three time periods were recognized (baseline period, intermediate period and follow up period) and modeled as statistical processes with respect to sensor observations, calculating their mean and control limits. Based on these control limits, days that do not lie within process control limits are candidate abnormal points.

Analysis of long-term sensor observations is presented in one case study, which refers to recognition of depressive symptoms' deterioration.

Participant #5.

Participant #5 is 71 years old and lives alone. She presents with symptoms of depression of which the most eminent are her lack of interest in activities and her frequently expressed sadness. Her mood fluctuates throughout the day from happy and energetic to pretty sad and tired. Loneliness and bad quality of sleep are important factors of her symptoms of depression. Other important factors are her poor capability to concentrate on activities and her fear of having memory losses. Also her mobility is limited because of her arthritis. Her knees are a source of severe and persistent pain which also affects negatively her mood.

Participant's clinical assessment of depressive symptomatology is provided for all three assessment periods: baseline, 1-month interim and 2-months follow up.

Table 1. Participant #5 depressive symptoms. Red cells indicate symptoms' deterioration.

	PHQ-1 (loss of interest)	PHQ-2 (depressive mood)	PHQ-7 (concentration deficits)
Baseline 27/1/2015	1	2	3
Interim 27/2/2015	3	2	1
Follow up 17/3/2015	3	3	3

Based on correlations that were found in the AHA LL data between sensors and diagnostic tools, concentration deficits severity is inversely proportional to number of steps, walking speed and speech arousal. Therefore, all three parameters are modeled and their statistical properties – the three parameters are modeled as statistical processes as described in [7] - are calculated for time periods where state deterioration is annotated according to PHQ-9 (c.f. **Table 1**). The whole monitoring period is divided in three time periods: the baseline period, - which accounts for a 2-week period, starting from the date that the baseline assessment was performed -, the interim period,- which accounts for the period starting right after the end of the baseline period and ending at the time of the interim visit and assessment was performed-, and the follow

up period, which accounts for the period starting right after the end of the interim period and ending at the time that follow up assessment is performed, at the end of the trial period.
Connected line represents the parameter's value fluctuation during the reference period, while the dots represent parameter's values during the period under investigation. Horizontal lines represent the statistical properties of the reference period, namely the mean process value, the lower and upper control limits (green, yellow and red color lines respectively). Values out of reference period's control limits may be considered as "abnormal" values and need to be interpreted according to the given context.

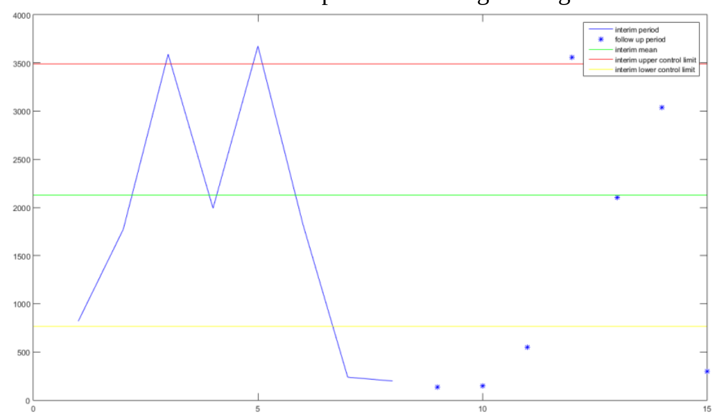

Fig. 5. Participant #5 step count modelling. Horizontal axis represents day number. Vertical axis represents the total number of steps per day.

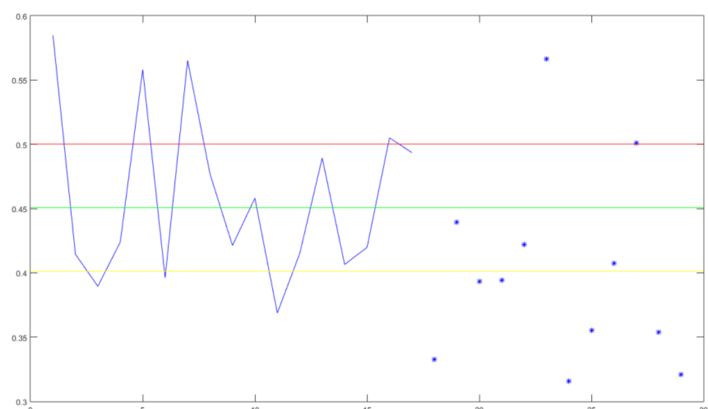

Fig. 6. Participant #5 walking speed modelling. Horizontal axis represents day number. Vertical axis represents daily average walking speed measured in meters/second (m/s).

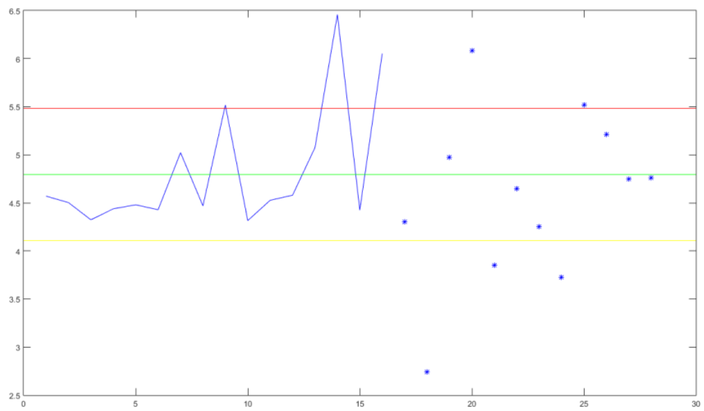

Fig. 7. Participant #5 speech arousal modelling. Horizontal axis represents day number. Vertical axis represents daily average speech arousal measured in abstract units (1-10).

All three figures show a decreasing trend in number of total steps per day, daily average walking speed and speech arousal. The decreasing trend is expressed in terms of follow up period days that lie below the lower control limit (yellow horizontal line) of the interim period. However, this is more apparent to the modeling representation of the mobility and gait parameters, rather than in the speech modeling. Concentration deficits of the participant seemed to got worsen according to the ground truth provided by the neuropsychological assessment. Therefore, there exists a correlation with the decreasing trends of the three parameters and the seniors' cognitive status.

4 Discussion

Three clinical scenarios were piloted in the AHA LL: monitoring of emotional disturbances, cognitive decline and functional ability. According to sensor analysis, quality of life and depressive symptomatology are related to mobility quantified as walking speed, step count and feet elevation. Through this kind of identification specific directions can be followed for both early diagnosis and accurate treatment. Elderly people quality of life is strongly related to physical performance [12], and therefore, there is a need to early detect any decreasing trends.

Robust measurement of health parameters in ecologically valid environments is a very important step, towards integration of intelligent monitoring systems in seniors' homes. We need to stress the fact that the protocol of activities that was used in the lab pilots, led seniors to behave in a free-form manner, being themselves and not having the belief and the anxiety they were assessed or monitored. This fact strengthens the results that have been obtained and is obviously along the lines of the overall system objectives which call for unobtrusiveness.

Pilots at home focused on the potential of using the technology developed within the project in real-life settings and provide evidence regarding its efficacy as a daily assistive tool for the elderly. Long term monitoring of seniors, based on lab evidence allow for deriving safer conclusions about the intelligent monitoring aspects of the monitoring system. Trend analysis presented preliminary evidence on decreasing health patterns, as sensor measurements were tested against changes annotated by neuropsychologists with clinical assessment tests. However, a two month monitoring period is considered as a limitation of our study, since it does not allow to check for slow varying disease trends, such as cognitive decline. This way the reason that just one case study was presented, since no significant health changes were observed to the rest of the participants, during the two-month period. However, since equipment is already in place in a limited number of homes, we plan -for those individuals that will accept the system to continue to be in their homes, - to allow for its existence for another six months or 1 year period. In this way, more validated data may be gathered and multiple follow up measurements may be obtained. The latter will provide useful insights not only for the health and quality of life of the involved individuals but for the entire health care system per se.

Acknowledgements.

Part of the research leading to these results has received funding from the European Union's Seventh Framework Programme (FP7/2007-2013) under grant agreement no 288532. For more details, please see http://www.usefil.eu. A.S. Billis is supported by a scholarship from Fanourakis Foundation (http://www.fanourakisfoundation.org/).

References

1. M. Popescu, G. Chronis, R. Ohol, M. Skubic, and M. Rantz, "An eldercare electronic health record system for predictive health assessment.", In e-Health Networking Applications and Services (Healthcom), 2011 13th IEEE International Conference on, pp. 193-196, 2011.
2. N. Noury, T. Hadidi, M. Laila, A. Fleury, C. Villemazet, V. Rialle, and A. Franco, "Level of activity, night and day alternation, and well being measured in a smart hospital suite", Proc. 30th Annual Int. Conference of the IEEE-EMBS, pp. 3328–3331, 2008.
3. Riboni, D., Bettini, C., Civitarese, G., Janjua, Z. H., & Bulgari, V. (2015). From lab to life: Fine-grained behavior monitoring in the elderly's home. In Proc. of PerCom Workshops. IEEE Comp. Soc.
4. Alemdar, H., Tunca, C., & Ersoy, C. (2014). Daily life behaviour monitoring for health assessment using machine learning: bridging the gap between domains. Personal and Ubiquitous Computing, 1-13.
5. Cheng, H., Liu, Z., Zhao, Y., Ye, G., & Sun, X. (2014). Real world activity summary for senior home monitoring. Multimedia Tools and Applications, 70(1), 177-197.
6. Artikis, A., Bamidis, P.D., Billis, A., Bratsas, C., Frantzidis, C., Karkaletsis, V., Klados, M., Konstantinidis, E., Konstantopoulos, S., Kosmopoulos, D., Papadopoulos, H., Perantonis, S., Petridis, S. & Spyropoulos, C.S., (2012, August). Supporting tele-health and AI-based clinical decision making with sensor data fusion and semantic interpretation: The USEFIL case study. Paper presented at the 1st International Workshop on Artificial Intelligence and NetMedicine, Montpellier, France.
7. Billis AS, Papageorgiou EI, Frantzidis CA, Tsatali MS, Tsolaki AC, Bamidis PD. A Decision-Support Framework for Promoting Independent Living and Ageing Well. IEEE J Biomed Heal Informatics 2015;19:199–209.
8. Mackintosh, J. H., Kumar, R., & Kitamura, T. (1983). Blink rate in psychiatric illness. The British Journal of Psychiatry, 143(1), 55-57.

9. Drummond, P. D. (1994). The effect of anger and pleasure on facial blood flow. Australian Journal of Psychology, 46(2), 95-99.
10. Siegman, A. W., & Boyle, S. (1993). Voices of fear and anxiety and sadness and depression: the effects of speech rate and loudness on fear and anxiety and sadness and depression. Journal of Abnormal Psychology, 102(3), 430
11. Purser, J. L., Weinberger, M., Cohen, H. J., Pieper, C. F., Morey, M. C., Li, T., ... & Lapuerta, P. (2005). Walking speed predicts health status and hospital costs for frail elderly male veterans. *Journal of rehabilitation research and development*, *42*(4), 535.
12. Rejeski, W. J., & Mihalko, S. L. (2001). Physical activity and quality of life in older adults. The Journals of Gerontology Series A: Biological Sciences and Medical Scienc-es, 56(suppl 2), 23-35.