ICT4Depression

User-friendly ICT Tools to Enhance Self-Management and Effective Treatment of Depression in the EU

Grant Agreement Number: FP7 – 248778

Deliverable 5.1r2:

Pilot study to test technical setup and tune parameters

Due date of deliverable: 31-01-2012
Actual submission date: 29-02-2012

Start date of project: 01-01-2010 Duration: 36 months

Coordinator	VU University Amsterdam
Deliverable leading partner	VU University Amsterdam
Contributing partners (this deliverable only)	All partners involved in the project.
Revision	2

<table>
<tr><th colspan="3">Project co-funded by the European Commission under the Seventh Framework Programme (FP7)</th></tr>
<tr><th colspan="3">Dissemination Level</th></tr>
<tr><td>PU</td><td>Public</td><td>X</td></tr>
<tr><td>PP</td><td>Restricted to other programme participants (including the Commission Services).</td><td></td></tr>
<tr><td>RE</td><td>Restricted to a group specified by the consortium (including the Commission Services).</td><td></td></tr>
<tr><td>CO</td><td>Confidential, only for members of the consortium (including the Commission Services).</td><td></td></tr>
</table>

Contents

1. Introduction and method 4
2. Results.. 5
3. Conclusion 15

1. Introduction and method

The goal of task 5.1 was testing the feasibility of wearing the devices that were developed as part of the project (i.e. the physiological devices developed by PLUX in the form of a glove and a chest strap, and the usage of accelerometer data from the mobile phone) for a prolonged period of time and the ability to reliably detect the ongoing activity (specifically posture and physical activity) of patients using these sensors. This was tested in a standardized ambulatory study at the VU University in Amsterdam. For this study students were recruited through announcements at the VU University. Students were provided written information about the study and they were asked for their informed consent. Twenty-seven healthy subjects underwent a standardized protocol while wearing the devices. The subjects received a mobile phone, a glove and a chest strap. The mobile phone was used to measure posture and activity as it contains a mobility monitor. To measure accurately posture and activity the mobile phone had to be worn in the subjects' pocket of the trousers. The glove had two electrodes, one for measuring blood volume pulse rate (BVP) and one for measuring electrodermal activity (EDA) which is a measure for skin conductance. The chest rate strap was used to measure heart rate and respiration.

The protocol consisted of two parts; a 1-hour supervised protocol and a 23-hour unsupervised part. In the 1-hour supervised part of the protocol, subjects engaged in the following activities under supervision of the experimenter: 1) 1 minute of walking; 2) 2 minutes sitting, 3) 2 minutes standing, 4) 2 minutes lying, 5) 2 minutes sitting, 6) 2 minutes lying, 7) 2 minutes standing, 8) 2 minutes sitting, 9) tone avoidance task, 10) 3 minutes walking, 11) 3 minutes walking and talking, 12) 2 minutes climbing stairs, 13) 3 minutes sitting, 14) 3 minutes cycling at 50W, 15) 3 minutes cycling at 100W, 16) 3 minutes cycling at 150W, 17) 3 minutes sitting, 18) 3 minutes walking at 5 km/hour, 19) 3 minutes walking at 6 km/hour, 20) 3 minutes walking at 8 km/hour, and 21) 3 minutes sitting.

After this 1-hour supervised protocol, full unsupervised ambulatory recording is done for a 23- hour recording period (second part of the protocol). During this period the subjects wore the sensor devices. Each 30 minutes they received a small questionnaire on an iPod to assess: 1) how they felt (scoring on a 5 points Likert scale), 2) what they were doing (% sitting, % lying, % walking, % cycling), and 3) with who they were. The next morning the subjects gave back the devices.

In the next chapter, results are described based on the 1-hour supervised part of the protocol (par 2.1) followed by the results of the unsupervised part of the protocol (par 2.2).

2. Results

2.1 Supervised part

The results using the experimental setup described in Chapter 1 are twofold: (1) data has been obtained on the usability of the devices, and (2) the accuracy of detecting physical activity and posture could be evaluated. First, the usability is described after which the accuracy of the second part is reported for each of the measurement devices.

2.1.1 Usability

The study was also used to test the usability of wearing the glove and the chest strap for a prolonged period. With regard to the usability the following observations were made:

1. The connection between the electrodes of the glove and the skin was not very good. Therefore, the following two solutions were applied; The electrodes were replaced by pre-gelled electrodes and an extra strap was fastened on the glove to assure a better fit around the wrist.
2. The connection between the electrodes of the chest strap and the skin was not always very good. Especially the fit of the heart strap for the women resulted in a weak connection. Therefore, the heart strap was adapted during the experiment for the women; an extra strap was made over the shoulder. Also, the normal electrodes were replaced by pre-gelled ones.
3. Signals from the sensors in the glove were very sensitive for movement resulting in some noise in the data during high intensity activities.

As a result of above mentioned points, data was not available for all participants during the whole period of the 1-hour protocol for each of the measurement devices. So, results are based on available data, ranging from 3 till 16 participants.

2.1.2 Mobility monitor

The first scripted part of the protocol was used to verify the performance of the mobility monitor. To this end, users were requested to leave the phone in one of their trousers' pockets whilst performing the activities in the order dictated by the protocol. A researcher observed all activities and accurately recorded start and stop times of all the activities. These time data were consecutively used in Matlab to verify the performance of the mobility algorithms.

Table 1: Confusion matrix for experiment without initial walking activity

	Lying	Sitting	Standing	Walking	Cycling
Lying	99.7	0	0	0.2	0.1
Sitting	66.5	17.8	14.7	0.4	0.6
Standing	87.6	0	11.6	0.2	0.6
Walking	7.4	2.5	0.5	68.2	21.4
Cycling	0	14.0	0.8	42.1	43.1

Table 1 lists the average confusion matrix over 10 trial runs. The confusion matrix expresses how often a certain activity (rows of the matrix) was qualified correctly or incorrectly (with the correct classifications occurring on the main diagonal) and the following can be concluded:

1. The difference between static (lying, sitting and standing) and dynamic activities (walking and cycling) can be determined with high accuracy; 99.3% of static activities is correctly classified as being static and 87.4 % of dynamic activities is correctly classified as dynamic.
2. Lying is identified correctly in 99.7% of cases.
3. Sitting is misclassified as lying or standing in 81.2% of cases.
4. Standing is misclassified as lying or sitting in 84.3% of cases.
5. Walking is correctly classified in 68.2% of cases.
6. Cycling is correctly classified in 43.1% of cases.

The misclassification of the static activities as another static activity has two reasons:

1. The mobility algorithms establish a base line on start-up and require the user to perform an activity that can be recognised without knowing the orientation of the device and that can be used to consecutively infer the orientation of the device relative to the user. For this purpose walking is used as the cadence of walking can be accurately identified without knowing the device of the user. Through the assumption that the user is walking upright, the orientation of the device can then be estimated. This action was not performed in these trials and hence most static activities prior to the first walking activity were misclassified.
2. With the mobile phone in the user's pocket, there is a significant chance that the phone's orientation does not change between sitting and lying activities. Hence, these two activities could not be distinguished in most trial runs.

To investigate the influence of the calibration cycle through the identification of a walking period, six trial participants was requested to walk for 1 minute prior to the start of the protocol. The resulting confusion matrix is shown in Table 2.

Table 2: Confusion matrix for experiment with initial walking activity

	Lying	Sitting	Standing	Walking	Cycling
Lying	48.6	51.4	0	0	0
Sitting	38.6	49.6	10.7	0.5	0.6
Standing	0	0	99.5	0.3	0.2
Walking	0.2	0	0.4	70.0	29.4
Cycling	0	0.1	1.2	30.0	68.7

To compare the confusion matrix thus obtained to the results obtained without an initial walking activity:

1. The difference between static (lying, sitting and standing) and dynamic activities (walking and cycling) can be determined with high accuracy; 100% of static activities is correctly classified as being static and 99.85 % of dynamic activities is correctly classified as dynamic. This is a marked improvement.
2. Lying is identified correctly in 48.6% of cases which is significantly lower than the results obtained earlier. This can be explained by the structure of the mobility algorithm: the lying activity is identified before sitting, and will thus take precedence if the algorithms have not yet established the difference between lying and sitting in terms of the orientation of the device. The fact that the correct classification is still low, is due to the aforementioned small difference in device orientation between sitting and lying activities.
3. Sitting is misclassified as lying or standing in 49.3% of cases. Although still high, this is partly (in 38.6% of cases) due to the aforementioned small difference in device orientation in lying and sitting situations.
4. Standing is misclassified as lying or sitting in 0% of cases.
5. Walking is correctly classified in 70% of cases.
6. Cycling is correctly classified in 68.7% of cases

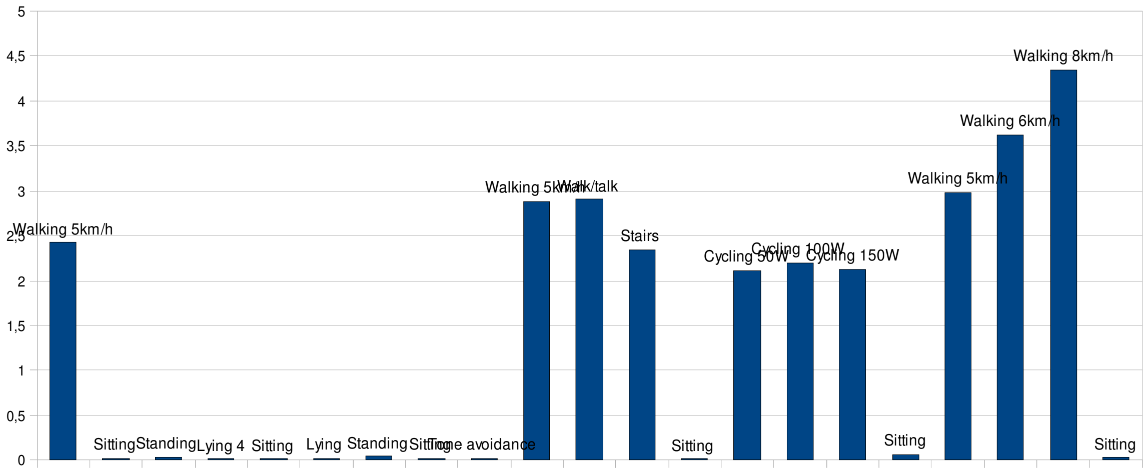

Figure 1: Mean values for energy expenditure, expressed in Counts per Minute (CPM) during all activities in the 1-hour protocol.

Comparing the confusion matrices in Table 1 and Table 2 shows that the addition of a walking activity resulted in a significant improvement in performance. A further analysis focussed on the verification of the information provided by the Counts per Minute (CPM) which is a measure for energy consumption derived from the acceleration of the user towards the earth (projection of acceleration vector onto gravity vector). Average results are shown in Figure 1. These results confirm the information presented in the confusion matrices and show that a clear distinction can be made between static and dynamic activities. Walking at various increasing speeds results in an increase in energy consumption recorded. This same relationship cannot be seen when analysing the cycling activities. This is due to the fact that increasing cycling activity was obtained through an increase of the cycling resistance, which for most users resulted in a lower rotational velocity, which in turn resulted in lower accelerations.

2.1.3 Chest strap

Two sensors are located in the chest strap, one for measuring respiratory frequencies (breathing rate) and one for measuring heart rate.

Figure 2 displays respiratory frequencies for all activities. Frequencies were quite steady during static activities as sitting, standing and lying. As expected, frequencies are higher during dynamic activities as walking, climbing stairs and cycling.

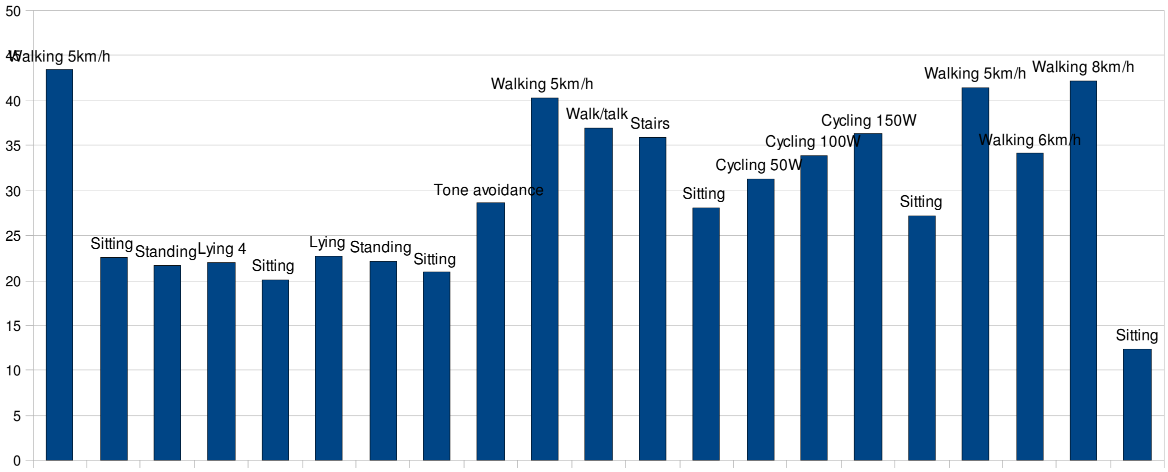

Figure 2: Mean values for respiratory frequency during all activities in the 1-hour protocol.

Figure 3 displays mean heart rate values for the activities. Heart rate is clearly higher during dynamic activities than during static activities. Exceptions are standing, the 3-minutes sitting after climbing stairs and cycling. It's not clear why the heart rate during standing was higher than during lying and sitting. The relative high heart rate during sitting after climbing stairs and cycling could be explained by the fact that participants are still tired after dynamic activities like climbing stairs and cycling resulting in a higher heart rate.

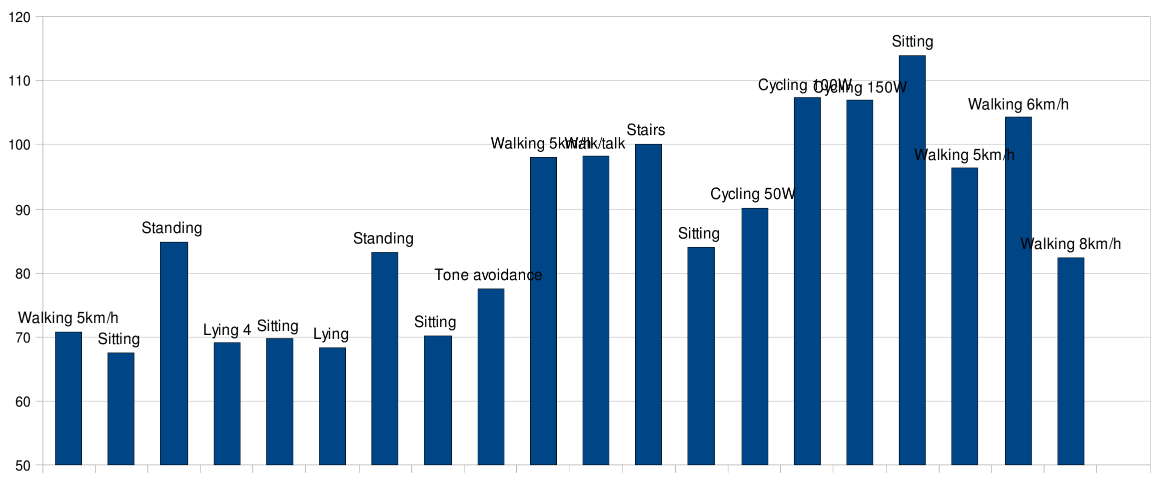

Figure 3: Mean values for heart rate during all activities in the 1-hour protocol.

2.1.4 Glove

The glove has two sensors, one for measuring blood volume pulse rate and one for measuring electrodermal activity (EDA).

Figure 4 shows the mean pulse rate during all activities. No clear relation is visible between pulse rate and the intensity of the activity. The values deviate also from the values of the mean heart rate while it was expected that the heart and pulse rate would show similar results.

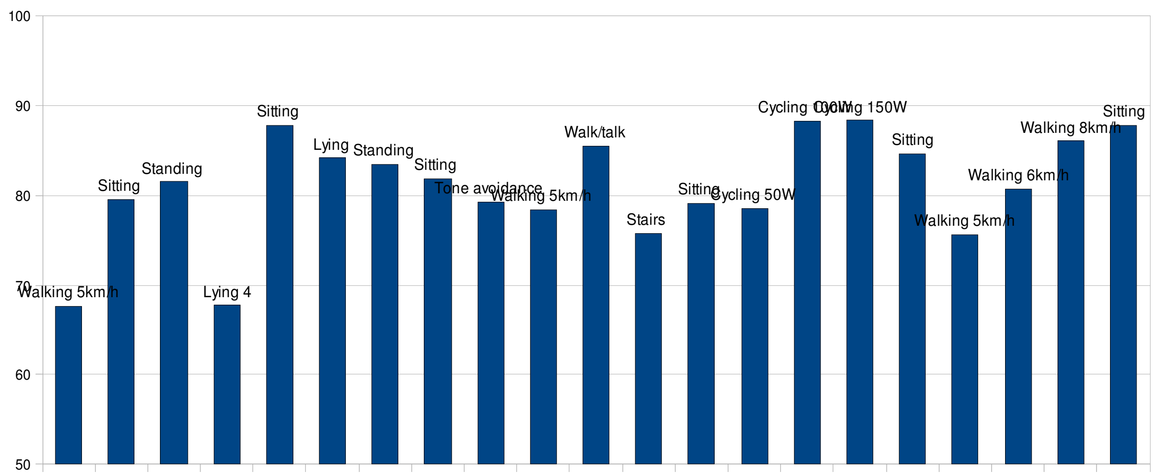

Figure 4: Mean values for blood volume pulse rate during all activities in the 1-hour protocol.

Correlations between heart rate and pulse rate were low (cycling 100W, r = -.02, lying, r =.30, walking 8km/h, r = -.52). The absence of a clear relation between heart rate and pulse rate is probably due to connection problems with the skin.

An EDA sensor is designed to measure skin conductance. The results of the magnitude of the electrodermal activity are shown in Figure 5. Results show higher electrodermal activity during dynamic activities as climbing stairs, cycling and walking in comparison to static activities as standing, lying and sitting. The relative high value during sitting after cycling can be explained by the fact that participants are still sweating because of the three cycling periods. The relative high value of electrodermal activity during the Tone avoidance task is striking. This task is meant to produce stress. And skin conductance is one of the fastest responding measures of the stress response. These results indicate that the sensor for measuring EDA is sensitive for detecting changes in skin conductance.

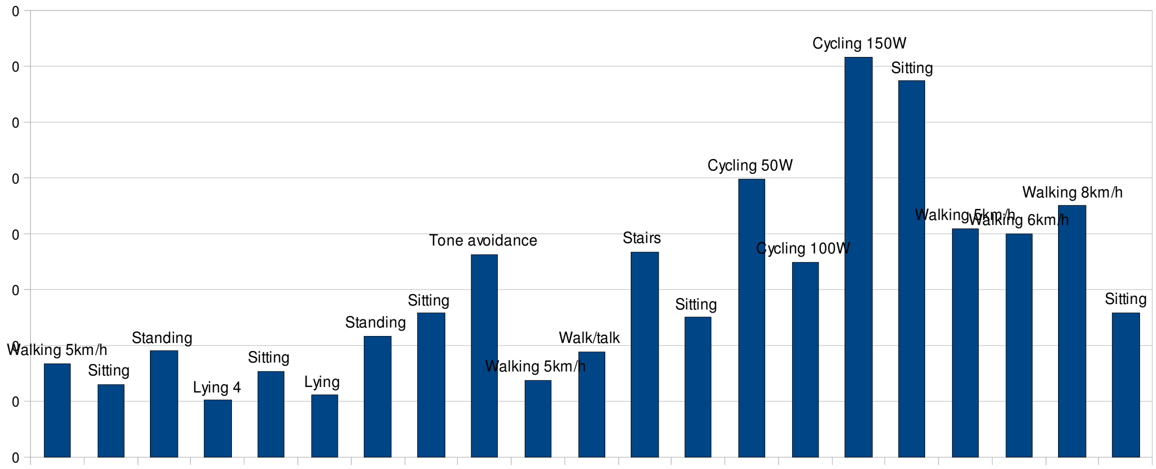

Figure 5: Mean values for electrodermal activity (amplitude) during all activities in the 1-hour protocol.

2.1.5 Conclusion

The activity monitor yields satisfactory results for the classification of static versus dynamic activities. Furthermore, the mobility monitor can be used to reasonably accurately detect activities such as walking and cycling, although further improvements are possible. These improvements can be derived from the addition of heart rate in the identification of activities: whereas the accelerometer data does not result in higher energy expenditure measurements with increasing cycling power (due to the fact that leg revolutions, and thus accelerometer output, decrease with increasing power), the increase in heart rate is evident and can be used to improve the predictions made by the accelerometer for energy expenditure which is very valuable for the project as both activities can be recognized better (to understand what the patient is doing) and also the desired increase in activities can be signaled better.

Frequencies in breathing rate and heart rate are clearly higher in dynamic activities compared to static activities. Pulse rate does not vary between enough between different activities. Skin conductance varies substantial between different activities, also during the tone avoidance task. Sitting activity and tone avoidance yield very clear differences in: heart rate, breathing rate and skin conductance. This clearly indicates that emotional markers can be identified by the PLUX sensors, which is certainly not possible with merely the accelerometer data. A further improvement in detection capabilities will be derived from the inclusion of chest orientation (through the accelerometer available in the PLUX chest sensor) to detect the difference between lying and sitting.

Results regarding pulse rate and skin conductance should be interpreted carefully as these data was sensitive for movement. Therefore, it's recommended to use the glove only during static activities in the open study.

2.2 Unsupervised part

Results of the unsupervised part of the experiment are based on 17 participants. Of these participants, data is available from one or more sensor devices. During the experiment, it became obvious that the battery of the mobile phone needed to be charged regularly, that is, each 5 or 6 hours. Also, some problems with the chargers of the chest strap and the glove were encountered; chargers worked not optimal resulting in data loss. Therefore we decided that participants did not need to wear the sensor devices during the night, instead the night was used to charge all batteries. Problems with the chargers of the chest strap and the glove were fixed half way the experiment. Because of charging the sensor devices during the day and the connection problems (see page 5), the amount of available data varies considerably between participants.

In this section, the following research questions are answered: 1. What is the relation between the different physiological parameters? 2. What is the relation between activities/postures as measured by the mobile phone and self-reported activities?

2.2.1 Relation between physiological parameters

Around 28 Ipod measurements were taken between the time the participant left the Lab of the University and midnight. First, mean values of the different physiological parameters were calculated for each time period between the measurements of the Ipod. Then, correlation analyses were performed between the parameters for each of these periods. A within-subject design was used; correlations were calculated for each participant separately. Table 3 displays the correlations between the physiological parameters for each participant separately.

The amount of measurements varies between participants; from minimal 4 till maximal 26. Correlations between heart rate and pulse rate were mainly positive. High correlations (0.7-0.8) were expected, however, just in three cases significant positive correlations were found. Correlations between heart rate and respectively respiration frequencies and energy expenditure (counts per minute) were in the majority of the cases positive and in some cases significant. However, for some persons a negative relation was found between heart rate and respiratory frequency. Correlations in both directions were found between heart rate and electrodermal activity (skin conductance).

Table 3: Individual relations between physiological parameters for each participant.

		HR/BVP		HR/RESP		HR/CPM		HR/EDA
subject	**n**	**r**	**n**	**r**	**n**	**r**	**n**	**r**
1								
2	10	0.31	10	-0.4	10	-0.2	10	.46
3	21							
4	12	0.48	11	0.43	15	-0.06	11	.19
5	19	0.23	10	-0.04	19	**0.71**	19	.05
6	20	0.55	18	**0.48**	21	**0.57**	21	-.25
7	4	-0.24	4	0.22	4	0.09	4	**.97**
8	20	**0.69**	20	**0.75**	20	0.67	14	-.14
9	6	0.39	6	0.38	6	0.06	6	.33
10	17	0.29	6	0.57	17	0.33	15	**.64**
11	16	-0.1	16	-0.34	16	-0.3	16	**-.43**
12	14	0.39	14	0.39	14	**0.56**	13	**-.62**
13	4	-0.12	4	0.76	4	0.81	4	.14
14								
15	8	**0.69**	8	**0.97**	8	0.52	8	.74
16	8	**0.94**	8	**-0.72**	8	0.16	8	.60
17	7	0.28	3	**-0.99**	8	**0.93**	7	-.33

n = number of within-subjects measurements, r = correlation coefficient, HR = heart rate, BVP = blood volume pulse rate, Resp = respiration frequency, CPM = counts per minute. EDA = electrodermal activity amplitude. Values in bold are significant (p <= 0.10).

Based on within-subjects measurements, analysis over the all group showed significant positive relations between heart rate on the one hand and pulse rate, respiratory frequency, energy expenditure and electrodermal activity on the other hand. See table 4.

Table 4: Overall relations between physiological parameters.

	HR	BVP	RESP	CPM	EDA	mood
HR		**.47**	**.38**	**.30**	**.19**	**-.17**
BVP			**.30**	.03	.05	.08
RESP				.11	-.09	**-.20**
CPM					.03	.06
EDA						.01

2.2.2 Relation between physiological parameters and mood.

Table 5 displays the correlations between the physiological parameters and self-reported mood for each individual participant. In general, few significant relations were found between self-reported mood on the one hand and energy expenditure, heart rate and electrodermal activity on the other hand. Analysis based on the total sample showed a negative relation between heart rate/respiratory frequency and mood (Table 4).

Table 5: Individual relations between physiological parameters and mood for each participant.

		mood/CPM		**mood/HR**		**mood/EDA**
subject	**n**	**r**	**n**	**r**	**n**	**r**
1	22	-0,22			21	.02
2	10	0	10	0,34	10	**.56**
3		-0,25			19	.21
4	19	0,1	15	-0,12	13	-.17
5	19	0,21	19	**0,47**	19	.28
6	21	0,16	21	0,02	21	-.30
7	14	0,25	4	0,24	14	.14
8						
9	8	0,3	6	0,65	8	.50
10	17	-0,15	17	0,16	15	.12
11	20	0,29	16	-0,38	16	.14
12	14	0,39	14	-0,3	13	.32
13	8	-0,77	4	-0,7	4	-.38
14	26	0,16				
15	26	0,14	8	-0,62	24	-.01
16	20	**0,42**			19	.01
17	21	0,35	8	-0,48	7	.34

n = number of within-subjects measurements, r = correlation coefficient, HR = heart rate, CPM = counts per minute, EDA = electrodermal activity amplitude. Values in bold are significant (p <= 0.10).

2.2.3 Activities measured by mobile phone and self-reported activities.

We calculated correlations between activities as recognised by the mobile phone and self-reported activities through the Ipod. The first twelve self-reported measurements of each participant were used. As can be seen in Table 6, low correlation coefficients were found for the correspondence between sitting, lying and cycling as recognised by the phone and through self-report. For standing and walking the correlations tend to be higher and, in some cases, significant correlations were found.

Table 6: Relations between self-reported activities and activities measured by phone.

	T1	T2	T3	T4	T5	T6	T7	T8	T9	T10	T11	T12
Sit	-	-	-.16	.31	.31	.30	.22	-	-.20	-	.19	.00
Lie	-.35	.26	.13	.23	.22	.26	.30	.20	.03	.22	-.38	-.25
Stand	-	-	.00	.11	-.14	-.08	**.63**	**.56**	-.29	**.58**	.13	**.59**
Walk	-	-.05	**.66**	.17	.42	.08	**.68**	**.86**	.41	.01	-.16	-.18
Cycle	-	-.39	-.15	.26	-.17	-.12	-.09	-.08	-	-.10	.50	.41

To further explore the capability of the phone to correctly classify the observed activity, a distinction was made between dynamic activities (walking and cycling) and static activities (sitting, lying and standing). In total, 311 within-subjects measurements were available. A significant positive relation ($r = .22$) was found between dynamic activities as measured by phone and self-reported dynamic activities. The same relation was found for static activities as measured by phone and self-report.

The relation between energy expenditure (stated as Counts Per Minute) and the % of dynamic activities (measured by phone) was studied over all participants. A significant positive relation was found of . 29 between CPM and % dynamic activities. Between CPM and % static activities a significant negative correlation was found of -.29. The correlations between CPM and the specific activities were -0.32 for lying, .10 for sitting, .13 for standing, .33 for walking and .22 for cycling.

3. Conclusion

During the supervised part of the experiment, the activity monitor yields satisfactory results for the classification of static versus dynamic activities. Furthermore, differences in physiological parameters were visible for different activities and postures. In the unsupervised part, positive relations were found between physiological parameters like heart rate, pulse rate, respiratory frequency, and energy expenditure on a group level. Also, the type of activity as detected by the phone corresponded with self-reported activities. However, the phone cannot always distinguish between static activities as lying, sitting and standing. When the user is static several issues occur: If the phone doesn't know its own orientation it cannot distinguish between lying, sitting and standing. If the phone is loosely in the pocket there will be no change in orientation between sitting and lying. The phone can be taken out of the pocket and can be lying on a table, which further skews results. The phone alone is not good at distinguishing sitting from lying (and standing in some cases) but we can use the PLUX accelerometer for this. Furthermore, results are skewed as undoubtedly the phone was not in the pocket for significant periods (i.e. when people were sleeping) .

On an individual level, the results vary considerably between participants. This is mainly due to the problems we encountered during the experiment, like connection problems between the devices and the skin and low battery duration. During the project, the wearable biomedical sensors were subjected to tests in terms of signal quality and usability. Results have shown the need for adjustments both in the glove and in the chest strap, and allowed us to characterize the situations in which each of the devices can operate reliably.

Concerning the chest strap, despite motion artifacts masking the signals, the problems were overcome using new digital filters. In terms of usability, during the activity of the year, several adjustments were performed in order to improve the textile to each person and the hygiene of the form-factors. Thus, three sizes were developed, and the colour of the textiles as well as the type were changed. The form-factor sizes allow the adaptation of the devices to each person, making them more comfortable for regular use by the patient. The colour made the textile materials more discrete and the type of fabric improves the everyday use by making them washable.

Based on this study, it is suggested to use data from the chest strap and glove sensor devices in the final pilot study in 2012 on an experimental basis, that is, using the data afterwards to make prediction models about symptom development.