Utility of Actimetry to Detect Apathy in Old-Age Depression: a Pilot Study

Manuel Abbas
Univ Rennes, Inserm, LTSI - UMR 1099
F-35000 Rennes, France
manuel.abbas@univ-rennes1.fr

Jean-Charles Roy
Academic Psychiatry Department, CHGR
Univ Rennes, F-35000 Rennes, France
jean-charles.roy@chu-rennes.fr

Gabriel Robert
Academic Psychiatry Department, CHGR
Inserm, IRISA UMR 6074, Empenn ERL U-1228
Univ Rennes, F-35000 Rennes, France
gabriel.hadrien.robert@gmail.com

Régine Le Bouquin Jeannès
Univ Rennes, Inserm, LTSI - UMR 1099
F-35000 Rennes, France
regine.le-bouquin-jeannes@univ-rennes1.fr

***Abstract*—Apathy is a promising marker of dementia conversion in old-age depression. However, the classical definition of apathy, as a lack of motivation, does not satisfy the criteria of objectivity and validity. The direct measure of subject activity by actigraphy might provide more promising results to characterize patients. In this paper, we evaluate the utility of actimetry to distinguish the activity of apathetic depressed elderly from non-apathetic depressed ones and control population. To this end, six features were extracted to characterize daily activity measured by a wrist-worn actimeter. These features present a statistical significance to discriminate between different groups. Furthermore, the results showed a linear relationship between the physical activity and the clinical severity of apathy. This pilot study suggests that the combination of motion signals parameters provides a promising diagnostic utility in apathetic depressed elderly, with an objective to complete current data with a larger cohort of participants.**

***Index Terms*—apathy, actimetry, feature extraction, detection, regression**

I. Introduction

Depression affects 7% of elderly population over 60 years [1]. In old-age depression, persistent depression symptoms account for the largest part of the increased risk of major cognitive disorders [2], [3], wherein apathy lies at the cornerstone [4]. Not only is it a frequently reported residual symptom of depression [5], apathy is also the most frequent behavioral and psychological symptom that precedes the onset of major cognitive disorders in various diseases [6], [7]. Hence, apathy is a promising marker of dementia conversion in old-age depression.

Classically, apathy is defined as a lack of motivation. But for a quantifiable trait of a given disorder to be considered as a diagnostic or prognostic marker, its measurement has to be objective, robust and reproducible. The psychometric evaluations of apathy, relying on the classical definition of apathy, do not satisfy these criteria, being dependent on the definition of motivation chosen by the authors [8] and limited by the patients' introspection abilities. A more objective approach would be to operationalize apathy as a reduction in goal-oriented behaviors [9]. Hence, apathy would become measurable by one's self-produced activity. Actigraphy is a non-invasive method employing an accelerometer which records one's minute-to-minute daily activity. Actigraphy studies have provided insightful results of ecological activity in depression [10], [11] as in apathy in elderly population [12]–[14]. Interestingly, it has been found that the combination of activity parameters [15] and the use of machine learning methods [16] had much more discriminatory properties between psychiatric states than isolated parameters alone.

In this study, we evaluate the utility of actimetry to distinguish the activity of apathetic depressed older adults from non-apathetic depressed ones and control population. To this end, we extracted six features from daily participants activity, and then evaluated their statistical significance using Kruskal-Wallis test as well as their ability to separate apathetic subjects from the rest using 1-*vs*-all and 1-*vs*-1 strategies as explained in the following sections. Moreover, the relationship between the characterized activity and the severity of apathy was assessed. This pilot study has an objective to reveal the existence or not of an underlying association between apathy and actimetry, with a view to an ongoing larger study. The remainder of the paper is organized as follows. Section II introduces the employed actimeter as well as the pre-processing and feature extraction operations. Section III presents the study cohort then illustrates and discusses experimental results. Finally, section IV discusses some limitations and suggests future considerations before concluding the paper in section V.

II. Materials and Methods

A. Raw Data & Pre-processing

Data were collected using a wrist-worn activity monitor, which is used to record information on the physical activity, namely ActiGraph wGT3X-BT. The Bluetooth® Smart

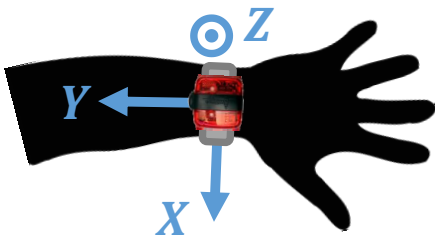

Fig. 1. Wrist-worn ActiGraph wGT3X-BT and the tri-axial orientation.

wGT3X-BT module integrates technologies around a tri-axial accelerometer and digital filtering technology. This device was worn around the wrist as illustrated in Fig. 1. ActiGraph uses an unit of measurement called 'count' in relation to activity measurements. Counts are the result of summing post-filtered acceleration values (using a band-pass filter, namely a 7th order IIR filter) around each axis into epochs, as indicated in [17], raw data being sampled at 30 Hz. Once the acquisition was done, acquired data were downloaded using ActiLife software. The epoch length represents the amount of time over which raw acceleration data are summed after filtering. Hence, a file consisting of counts values over the three axes is generated for each participant. Afterwards, data were segmented using 1-hour segments. Consequently, a $M \times N \times 3$ matrix resulted from previous operations, where M is the number of 1-hour segments (time-series), each consisting of N data-points for each axis $\{X, Y, Z\}$ and corresponding to a participant.

The time-series were analyzed to extract some parameters that might reveal the signs of apathy in terms of physical motion. If c_λ corresponds to the counts relative to the λ axis, $\lambda \in \{X, Y, Z\}$, the vector magnitude $||c||$ is computed as:

$$||c|| = \sqrt{c_X^2 + c_Y^2 + c_Z^2} \tag{1}$$

The time-series were also scaled using the z-score technique, so that the mean and the standard deviation (SD) of the resultant signal are equal to 0 and 1 respectively:

$$\hat{c}_\lambda = \frac{c_\lambda - \mu_\lambda}{\sigma_\lambda} \tag{2}$$

with μ_λ and σ_λ the mean and SD values of c_λ respectively.

B. Feature Extraction

Once the pre-processing was done, we proceeded to feature extraction, leading to six features detailed hereafter.

Feature F_1: the activity level is measured every 200 seconds (orange crosses in Fig. 2). F_1 is the mean of these levels. This feature takes into consideration (i) inactivity periods and (ii) intensity of movements. It outperforms traditional features like SD and range of signals computed globally. This feature was calculated over $\{\hat{c}_X, \hat{c}_Y, \hat{c}_Z\}$ components.

Feature F_2: this measure expresses local time-series forecasting. It provides local median prediction δ using the past $m = 3$ values throughout the time-series:

$$\begin{cases} \delta_\lambda = [c_\lambda(i), ..., c_\lambda(i+m-1)] \\ e(i) = \overline{\delta_\lambda} - c_\lambda(i+m) \end{cases} \tag{3}$$

with $c_\lambda(i)$ the i^{th} element of c_λ, $\overline{\delta_\lambda}$ the median of δ_λ and e the residuals.
To encode this information into a feature, e is divided into 5 equal segments without overlapping, and the mean of each of these segments is calculated and stored in an array E. Hence:

$$F_2 = \frac{\sigma_E}{\sigma_e} \tag{4}$$

This feature was calculated over $\{\hat{c}_X, \hat{c}_Y, \hat{c}_Z\}$.

Feature F_3: this feature quantifies, in a certain way, the trend in $||\hat{c}||$, consisting of N data-points, by calculating the mean of its cumulative sum Q:

$$\begin{cases} Q(t) = \sum_{i=1}^{t} ||\hat{c}(i)|| \\ F_3 = \frac{1}{N} \sum_{t=1}^{N} Q(t) \end{cases} \tag{5}$$

Feature F_4: the time-series were divided into 200-second fragments. The SD value of each fragment, revealing the intensity of movements, is calculated and stored in an array W. F_4 is the distribution entropy of W. This feature quantifies the randomness of the subject's movements (whether they are regular or not). Here, this feature was calculated over raw counts data $\{c_X, c_Y, c_Z\}$.

Feature F_5: to represent and analyze abrupt changes efficiently in time-series, wavelets are needed since they are well localized in time and frequency. Therefore, the 4-level wavelet decomposition of $\{\hat{c}_X, \hat{c}_Y, \hat{c}_Z\}$ is computed, using the 3rd-order Daubechies wavelet. Feature F_5 is the SD of the last level detail coefficients, evaluating their variability.

Feature F_6: the goal here is to fit an AutoRegressive-Moving-Average (ARMA) model using the first half of the

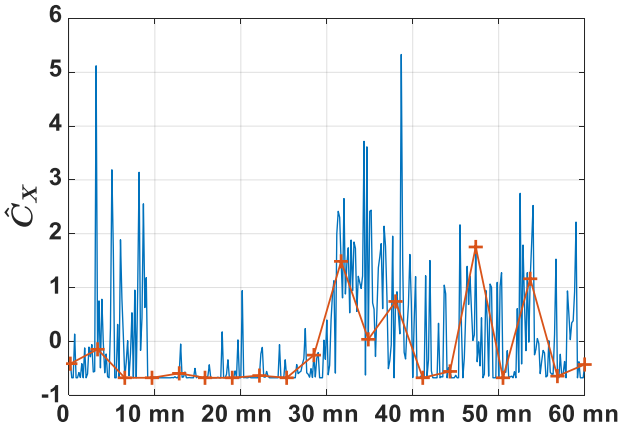

Fig. 2. An example of activity level (orange crosses) measured every 200 seconds over the x-component.

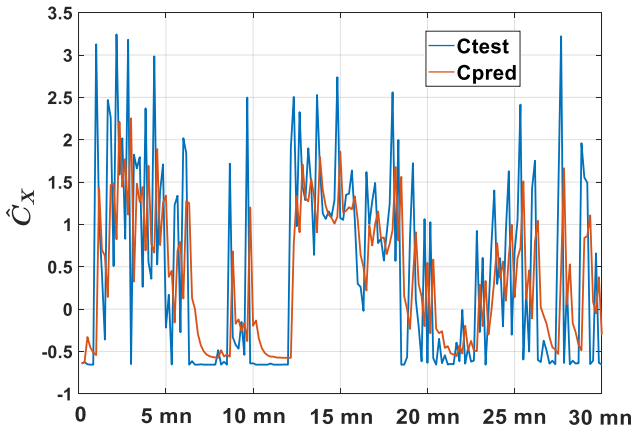

Fig. 3. Blue signal representing real counts values (Ctest) of the x-axis *vs* orange signal representing the output of the ARMA model (Cpred).

window (first 30 minutes). The equation of an ARMA model is given by:

$$c_\lambda(t) = \alpha + \epsilon(t) + \sum_{i=1}^{p} \phi_i X(t-i) + \sum_{i=1}^{q} \theta_i \epsilon(t-i) \tag{6}$$

with p the order of AR polynomial and ϕ its parameters, q the order of MA polynomial and θ its parameters, ϵ a white noise, and α a constant.
Afterwards, the step ahead prediction is computed using the second half of the window, *i.e.* predict the output of the fitted model using input-output data history from the second half of the window, as shown in Fig. 3. F_6 is the mean of the residuals, which are the difference between real data (Ctest) and the predicted data (Cpred). This feature encodes the information resulting from the representation of counts as a time-varying process, and was calculated over $\{\hat{c}_X, \hat{c}_Y, \hat{c}_Z\}$ with $p = 3$ and $q = 1$.

Consequently, 16 measurements are extracted in total from each sample, based on this feature extraction process.

III. Experimental Results

A. Study Cohort

Thirty-seven older adults (10 males and 27 females) of diverse profiles, whose age ranges between 67 and 87 years old (75.62 ±5.26), were recruited to conduct the study. They provided written informed consent prior to the experiments. Twenty-six subjects were clinically depressed and the severity of apathy was assessed with the Apathy Evaluation Scale (AES). The resultant scores range between 27 and 55. Among these 26 subjects, 7 were assessed as non-apathetic, while the other 19 subjects were considered as apathetic. The remaining 11 participants were non-depressed and therefore declared as control subjects. Consequently, they did not have any apathy score. Data were collected from participants over a period of 3 ±0.23 days using the wrist-worn ActiGraph wGT3X-BT during their daily routine. Data segmentation into 1-hour segments led to 2549 time-series per component. The epoch length was configured at 10 seconds, leading to $N = 360$ data-points in each 1-hour time-series. Hence, the resultant matrix is $2549 \times 360 \times 3$. The study was approved by the local ethics committee of Rennes University Hospital and conducted in accordance with the current French legislation.

B. Recognizing apathetic individuals

The efficacy of the aforementioned features in detecting apathy from motion signals was studied, in order to assess their capability to separate apathetic subjects from other individuals. In the remainder of this paper, apathetic people are denoted by 'A', depressed but non-apathetic individuals by 'NA', and control subjects by 'C'. Two strategies were considered to conduct the study, namely 1-*vs*-all (*i.e.* apathetic subjects *vs* remaining individuals) and 1-*vs*-1 (*i.e.* apathetic *vs* non-apathetic or control). Consequently, the features were averaged by subject, resulting in a 37×16 matrix, where each row corresponds to a subject and represents the average values per hour of the features. Furthermore, the values of each measurement, representing a column in the 37×16 matrix, are scaled using the following robust-outlier sigmoidal model [18], due to their heterogeneity in terms of magnitude, unit and range:

$$\hat{f}_j = \frac{1}{1 + \exp\left(-\frac{f_j - \overline{f}}{Q_f}\right)} \tag{7}$$

with $\hat{f}_j$ and f_j the j$^{\text{th}}$ element in the scaled feature vector and the original vector respectively, $\overline{f}$ the median of vector f, and Q_f the interquartile range (IQR) of f divided by 1.35. This non-linear model scales the values into [0, 1] and affects the distribution of data-points in the space.

Table I illustrates the median value and IQR of each scaled feature following the 1-*vs*-all strategy. Based on these two metrics, which give an idea about the boxplot of each population, we observe that the proposed features are globally able to separate both groups since the median values are relatively distant in the unity scale. To evaluate the statistical significance of these features, we ran the Kruskal-Wallis test to calculate the p-value. This test has the null hypothesis H_0 that individuals in each categorical group come from the same population. In general, the six features are statistically significant at a 5% significance level, except for F_1 over the y-component, representing activity levels following the arm axis, and F_4 & F_5 over the z-component, characterizing the "disorder" of movements intensity and the variability of high frequency components following this axis (up-down movements) respectively. Hence, they underline a certain relation between apathy and actimetry. The principal component analysis (PCA) was also applied on the scaled dataset to reduce the dimensionality by minimizing information loss, thus increasing the interpretability. It allows us to visualize data in 2D space, and thus to observe trends, jumps, clusters and outliers, and to see whether the two groups are separable or not. Fig. 4 illustrates the

TABLE I

THE EXPERIMENTAL RESULTS IN TERMS OF MEDIAN, INTERQUARTILE RANGE (IQR), AND P-VALUE FOLLOWING SIX FEATURES, USING 1-*vs*-ALL STRATEGY.

Feature	F_1			F_2			F_3	F_4			F_5			F_6		
	$\hat{c}_X$	$\hat{c}_Y$	$\hat{c}_Z$	$\hat{c}_X$	$\hat{c}_Y$	$\hat{c}_Z$	$\|\hat{c}\|$	c_X	c_Y	c_Z	$\hat{c}_X$	$\hat{c}_Y$	$\hat{c}_Z$	$\hat{c}_X$	$\hat{c}_Y$	$\hat{c}_Z$
Median $\binom{A}{O}^{\ddagger}$	$\binom{0.47}{0.63}$	$\binom{0.38}{0.54}$	$\binom{0.41}{0.6}$	$\binom{0.64}{0.42}$	$\binom{0.59}{0.44}$	$\binom{0.58}{0.45}$	$\binom{0.46}{0.61}$	$\binom{0.39}{0.54}$	$\binom{0.45}{0.59}$	$\binom{0.37}{0.6}$	$\binom{0.38}{0.63}$	$\binom{0.35}{0.56}$	$\binom{0.37}{0.55}$	$\binom{0.63}{0.4}$	$\binom{0.59}{0.42}$	$\binom{0.56}{0.44}$
IQR $\binom{A}{O}^{\ddagger}$	$\binom{0.29}{0.41}$	$\binom{0.4}{0.42}$	$\binom{0.29}{0.27}$	$\binom{0.33}{0.18}$	$\binom{0.27}{0.14}$	$\binom{0.38}{0.18}$	$\binom{0.25}{0.26}$	$\binom{0.3}{0.2}$	$\binom{0.16}{0.28}$	$\binom{0.38}{0.24}$	$\binom{0.24}{0.24}$	$\binom{0.3}{0.29}$	$\binom{0.28}{0.32}$	$\binom{0.32}{0.25}$	$\binom{0.26}{0.23}$	$\binom{0.25}{0.35}$
p-value†	0.005	> 0.05	0.01	0.04	0.01	0.04	0.03	0.04	0.04	> 0.05	0.002	0.02	> 0.05	0.01	0.005	0.02

‡A: Apathetic (top); O: Others (bottom) – †Kruskal-Wallis test

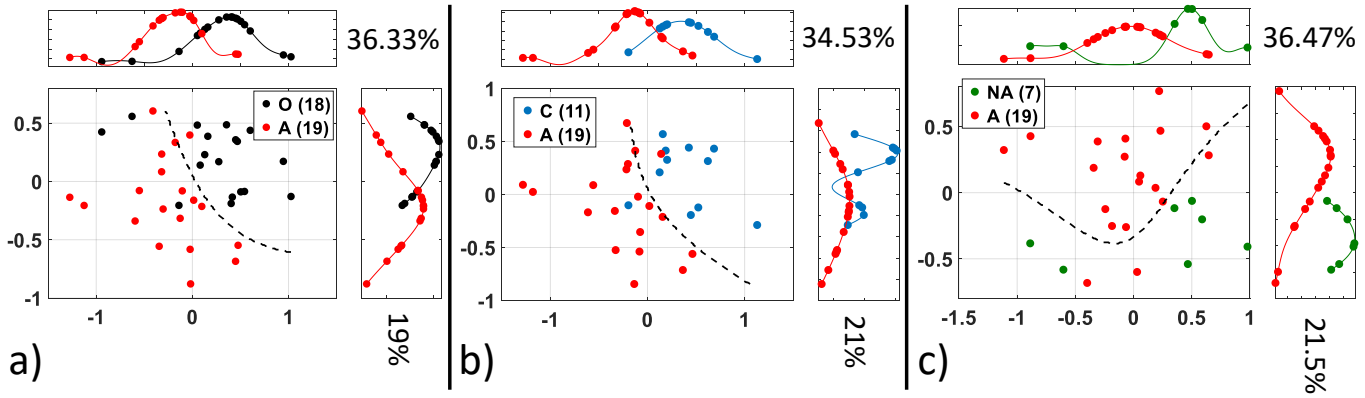

Fig. 4. Data distribution in 2D space after applying PCA, with the corresponding kernel smoothed densities (top and side distributions) for each pair of populations: (a) Apathetic *vs* Others, (b) Apathetic *vs* Control, and (c) Apathetic *vs* Non-Apathetic. The boundary of the SVM_{RBF} separating both populations is plotted in dashed line, and the percentage of variance explained by each principal component is indicated.

distribution of data-points following two principal components (PCs): (a) Apathetic *vs* Others (1-*vs*-all strategy), (b) Apathetic *vs* Control (1-*vs*-1), (c) Apathetic *vs* Non-Apathetic (1-*vs*-1). Each point represents a subject of the study cohort. A kernel smoothed density was estimated and plotted for each group using the scores of the PCs. These densities are represented by the top and side distributions. Additionally, Support Vector Machines (SVM) with different kernels have been tested to separate both groups in each case. The SVM with Radial Basis Function (RBF) kernel achieved the highest discrimination power, and its boundary is shown in the plots (dashed line). Moreover, the percentage of variance explained by each PC is written on the figure. Clearly, the apathetic population is distinguishable in the aforementioned three cases with the information provided by the proposed features. With the 1-*vs*-all strategy, one apathetic subject and three individuals belonging to the other group are on the opposite side of the SVM boundary (*i.e.* 1 false negative and 3 false positives). The variance of PC_1 and PC_2 are 36.33% and 19% respectively. An overlap exists between the top distributions as well as the side ones. When it comes to 'A' *vs* 'C', the overlap is relatively bigger between the side distributions. Nonetheless, the variance of PC_1 decreases but that of PC_2 increases. A false positive and a false negative are obtained. Finally, the third scenario ('A' *vs* 'NA') leads to a smaller overlap between distributions and a higher variance explained by PCs, without any false positive. However, three subjects are wrongly identified as 'NA'. This might be explained by the small number of 'NA' subjects compared to other population. It is worth mentioning that machine learning classification is not the purpose of this study, given the small number of samples. The main objective is to study the capacity to identify apathetic older adults, based on physical activity measured by actimetry. These preliminary results are quite promising, and show that both populations are non-linearly separable.

C. Estimating AES

The next goal is to estimate the AES score of 'A' and 'NA' subjects based on the PCs. Fig. 5 illustrates AES as a function of both PCs, showing that a linear relationship seems to exist except for few outliers. To this end, a linear regression 'R_L' model was tested, using Leave-Subject-Out (LSO) cross validation. In other words, the model was fitted to the observed data of 25 subjects (among the 26 depressed subjects) and tested on the remaining one at each iteration. To evaluate the LR model, we compared it to a quadratic SVM regression model denoted by 'R_Q' and a SVM_{RBF} regression model denoted by 'R_R'. Moreover, two other naive models were tested, the first one generating uniformly distributed random values between 27 and 55 denoted by 'R_U', the second one generating normally distributed random values using the mean and SD of AES scores and denoted by 'R_N'. The root-mean-squared error (RMSE) was calculated for these five models, which led to the following

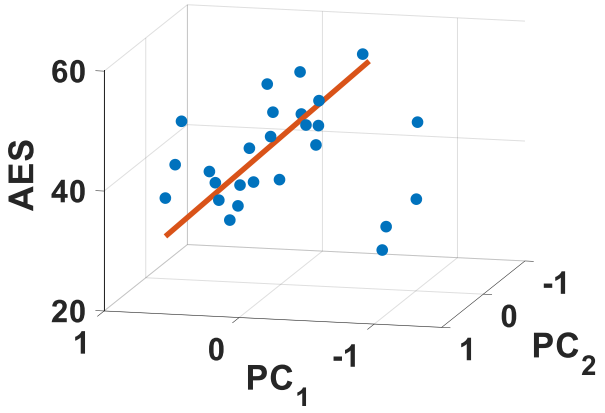

Fig. 5. Modeling the relationship between the two principal components and AES by fitting a linear equation of the form $AES = \alpha_1 PC_1 + \alpha_2 PC_2 + \beta$ after the exclusion of outliers. Each point represents a subject.

respective values:

$$\{R_L;\ R_Q;\ R_R;\ R_U;\ R_N\} \rightarrow \{9.1;\ 9.5;\ 10.1;\ 13.2;\ 12.8\}$$

The linear model outperforms the rest with an error of 9.1 for values ranging between 27 and 55. The existence of four distinguishable outliers makes the RMSE increase. The characteristics of the corresponding subjects (the outliers) will be investigated in a future work to try explaining their distance from the linear curve. In conclusion, the relationship between AES scores and the PCs is explained linearly, while data of 'A' and those of 'NA' following two PCs are non-linearly separable in the 2D space (see Fig. 4).

IV. Limitations and Future Considerations

The study cohort being relatively small, we are currently working on recruiting more older adults to generalize our results for classification on a larger dataset. Moreover, while the extracted features may show diagnostic or predictive utility, their interpretation from a clinical point of view would need further exploration on a larger scale. Furthermore, the analyzed data are counts, which are the result of summing post-filtered acceleration measures. This preprocessing is valuable to remove unwanted components before the analysis, but might also exclude some hidden information from raw acceleration data which may increase the discrimination power of the system.

A relatively long epoch length of 10 seconds is useful to decrease the computational complexity of feature extraction (fewer data-points in signals), but might discount some descriptive parts of signals that can be exploited. Therefore, studying the effect of epoch length on the results is an ongoing research topic. Moreover, the identification of long inactive periods and recurrent sedentary lifestyle is going to be an objective of a future work.

V. Conclusion

This paper investigates the utility of actimetry to detect apathy in old-age depression. Our pilot study showed that motion signals, measured by a wrist-worn sensor, might provide a good indicator of apathy. The proposed features were able to separate apathetic individuals from other subjects with a quite acceptable accuracy, after undergoing a non-linear transformation using the sigmoidal function and an orthogonal transformation by PCA, converting features into uncorrelated variables. We also found with these preliminary results that the linear regression model outperformed other models when it comes to the prediction of AES scores.

References

[1] World Health Organization. Mental health of older adults. Published: Dec. 2017. Available: https://www.who.int/news-room/fact-sheets/detail/mental-health-of-older-adults

[2] P.J. Brown, *el al.*, "The Depressed Frail Phenotype: The Clinical Manifestation of Increased Biological Aging," The American Journal of Geriatric Psychiatry, vol. 24, no. 11, pp. 1084–1094, June 2016.

[3] A.R. Kaup *et al.*, "Trajectories of Depressive Symptoms in Older Adults and Risk of Dementia," JAMA Psychiatry, vol. 73, no. 5, pp. 525–531, 2016.

[4] I. Groeneweg-Koolhoven *et al.*, "Apathy in early and late-life depression," Journal of Affective Disorders, vol. 223, pp. 76–81, Dec. 2017.

[5] J.R. Calabrese *et al.*, "Methodological approaches and magnitude of the clinical unmet need associated with amotivation in mood disorders," Journal of Affective Disorders, vol. 168, pp. 439–451, 2014.

[6] Z. Ismail *et al.*, "Neuropsychiatric symptoms as early manifestations of emergent dementia: Provisional diagnostic criteria for mild behavioral impairment," Alzheimer's & Dementia: The Journal of the Alzheimer's Association, vol. 12, no. 2, pp. 195–202, June 2015.

[7] K. Dujardin, P. Sockeel, M. Delliaux, A. Destée, and L. Defebvre, "Apathy may herald cognitive decline and dementia in Parkinson's disease," Movement Disorders: Official Journal of the Movement Disorder Society, vol. 24, no. 16, pp. 2391–2397, Dec. 2009.

[8] R.S. Marin, "Apathy: A neuropsychiatric syndrome," The Journal of Neuropsychiatry and Clinical Neurosciences, vol. 3, no. 3, pp. 243–254, 1991.

[9] R. Levy and B. Dubois, "Apathy and the functional anatomy of the prefrontal cortex-basal ganglia circuits," Cerebral Cortex, vol. 16, no. 7, pp. 916–928, Jul. 2006.

[10] A. Gershon, N. Ram, S.L. Johnson, A.G. Harvey, and M.J. Zeitzer, "Daily Actigraphy Profiles Distinguish Depressive and Interepisode States in Bipolar Disorder," Clinical psychological science, vol. 4, no. 4, pp. 641–650, Nov. 2015.

[11] A.C. Volkers, J.H.M. Tulen, W.W. van den Broek, J.A. Bruijn, J. Passchier, and L. Pepplinkhuizen, "Motor activity and autonomic cardiac functioning in major depressive disorder," Journal of Affective Disorders, vol. 76, no. 1-3, pp. 23–30, Sept. 2003.

[12] R. David *et al.*, "Ambulatory actigraphy correlates with apathy in mild Alzheimer's disease," Dementia, vol. 9, no. 4, pp. 509–516, Sept. 2010.

[13] A.M. Goldfine, B. Dehbandi, J.M. Kennedy, B. Sabot, C. Semper, and D. Putrino, "Quantifying post-stroke apathy with actimeters," The Journal of neuropsychiatry and clinical neurosciences, vol. 28, no. 3, pp. 199–204, Feb. 2016.

[14] J.M. Zeitzer *et al.*, "Daily Patterns of Accelerometer Activity Predict Changes in Sleep, Cognition, and Mortality in Older Men," The Journals of Gerontology. Series A, vol. 73, no. 5, pp. 682–687, 2018.

[15] J. Scott, A.E. Vaaler, O.B. Fasmer, G. Morken, and K. Krane-Gartiser, "A pilot study to determine whether combinations of objectively measured activity parameters can be used to differentiate between mixed states, mania, and bipolar depression," International journal of bipolar disorders, vol. 5, art. no. 5, 2017.

[16] P. Jakobsen *et al.*, "Applying machine learning in motor activity time series of depressed bipolar and unipolar patients compared to healthy controls," PLoS ONE, vol. 15, no. 8, e0231995, Aug. 2020.

[17] Actigraph. What are counts? Published: Nov. 2018. Available: https://actigraphcorp.my.site.com/support/s/article/What-are-counts

[18] M. Abbas, M. Saleh and R. Le Bouquin Jeannès, "A Hybrid Solution for Human Activity Recognition: Application to Wrist-Worn Accelerometry," Fifth International Conference on Advances in Biomedical Engineering (ICABME), 2019, pp. 1-4.