

DESARROLLO DE UN SOFTWARE LIBRE ORIGINAL PARA EL ANÁLISIS DE DATOS DE ACCELEROMETRÍA

DEVELOPMENT OF AN ORIGINAL SOFTWARE FOR
ACCELEROMETRY DATA PROCESSING AND ANALYSIS

Autor

Jorge Marín Puyalto

Tutores

José Antonio Casajús Mallén – Educación Física y Deportiva

Jorge Luis Ojeda Cabrera – Estadística e Investigación Operativa

Fecha de presentación

16/06/2017

Resumen

La práctica de actividad física y el sedentarismo están relacionados con el estado de salud general. Uno de los métodos más empleados para la evaluación objetiva de la práctica de actividad física cotidiana es el registro de acelerometría. Sin embargo, los programas informáticos disponibles para el análisis de estos registros tienen una escasa posibilidad de personalización y presentan dificultades para diferenciar el sedentarismo respecto del tiempo no válido. Por lo tanto, el objetivo principal de este estudio es el de desarrollar una herramienta informática para el procesamiento y análisis de registros de acelerometría de una forma rápida, versátil y accesible para el usuario. Además, se estudiaron los métodos y umbrales para la detección del tiempo no válido. El programa se desarrolló en R, un lenguaje de programación diseñado específicamente para el análisis estadístico. Para evaluar su capacidad, se utilizaron los registros de 35 participantes. Los participantes registraron el tiempo en el que no llevaron puesto el acelerómetro. La detección del tiempo no válido basada en la variabilidad y el rango de las aceleraciones se evaluó analizando la función de la distribución de probabilidad y las curvas receptor-operador. La versión final del código cuenta con 21 funciones y una plantilla markdown con un total de 765 líneas de código sin comentar. El software recibe datos de acelerometría y tras procesarlos y analizarlos, genera informes individuales y bases de datos grupales. 14 parámetros distintos pueden ser fácilmente modificados a través de una ventana. Se pueden analizar varios archivos con una única llamada a la función. Además, se proponen nuevos puntos de corte para la detección del tiempo no válido, basados en las características del acelerómetro. Esta herramienta gratuita permite una evaluación de la práctica de actividad física y los comportamientos sedentarios a través de la acelerometría de una manera rápida y sencilla.

Abstract

Physical activity participation and sedentary patterns are closely linked to overall health. Accelerometry registers are one of the most widely used methods to objectively evaluate daily physical activity. However, the currently available software to analyze raw accelerometer data has a limited capacity to be adapted to different research needs and shows deficiencies differentiating sedentary time with non-wear intervals. Therefore, the main goal of the present study is to develop a customizable tool to process and analyze accelerometry records in quick and user-friendly way. Additionally, the methods and thresholds to detect non-wear periods were examined. The software was developed in R, an open-source coding language specifically designed for statistical computing, and tested with accelerometry registers from 35 volunteers. Non-wear data was classified according to annotations from the participants. To establish the applicability of the analysis of acceleration range and variability for the detection of non-wear periods, density distribution functions and receiver-operator characteristics curves were constructed. The final version of the code is formed by 21 functions and a markdown template file totaling 765 lines of uncommented code. It takes as an input a raw accelerometer data file and, after a two-step process of formatting and analyzing the data, returns the processed data, individual reports and general databases. 14 different parameters can be customized through a prompt window. Several files can be consecutively analyzed with a single function call. In addition, the proposed cut-off points for non-wear detection, based on the measuring device characteristics, classify precisely the non-wear intervals. This free tool will allow a quick, easy and versatile evaluation of physical activity participation and sedentary habits from raw accelerometer data.

MeSH Keywords: Software design; Computer-assisted signal processing; User-computer interface; Physical activity; Sedentary lifestyles

Índice

1. Introduction	5
2. Methods	6
3. Results	7
3.1. Software.....	7
3.1.1. Data processing functions.....	9
3.1.2. Data analysis functions.....	9
3.1.3. Batch file processing.....	10
3.1.4. Individual report generation	11
3.1.5. Database creation	12
3.1.6. Graphic user interface	12
3.2. Non-wear detection	12
4. Discussion.....	14
4.1. Software development	14
4.1.1. Main features.....	14
4.1.2. Applications.....	15
4.2. Non-wear detection	16
5. Strengths and limitations of the study	16
6. Conclusions	17
7. References	17

Listado de abreviaturas

AUC: Area Under the Curve

ENMO: Euclidean Norm Minus One

GUI: Graphical User Interface

ROC: Receiver Operator Characteristics

1. Introduction

Physical activity participation is an important marker of health (1). A physically active lifestyle has been associated with a reduction in the risk of suffering different diseases (2). Several public health organizations, such as the World Health Organization or the U.S. Department of Health and Human Services, have issued recommendations of physical activity participation (3,4). Therefore, an accurate assessment of physical activity is necessary to evaluate overall health and can help detecting subjects with an increased risk of suffering an adverse event.

Several methods have been designed to assess physical activity. Energy expenditure during different activities can be precisely evaluated with indirect calorimetry (5). Doubly-labelled water is useful to assess energy expenditure over a time period (6). These methods are precise but unable to distinguish different exercise intensities during the test. Numerous physical activity participation questionnaires have been developed and validated for different populations (7–9). Their precision, however is limited due to the subjective nature of the method (10). Accelerometry registers represent a method to measure daily physical activity precisely and objectively (11–15).

Accelerometry equipment has evolved from bulky devices with limited battery and data storage with a single sensor to easily-wearable triaxial accelerometers that are able to measure and store data during long periods of time. Nowadays, one of the most used models is the GENEActiv, a water-resistant device acceleration, temperature and luminosity sensors that can be wrist-worn, validated for the measurement of human physical activity (15). These improvements in autonomy, sensor precision and user comfort have made accelerometry a commonly used method to objectively assess habitual physical activity.

However, the high volume of data generated during an accelerometry register requires computer processing in order to extract results. The official software for the GENEActive devices (available from <https://open.geneactiv.org/index.html>) is very limited in terms of analysis customization and presents errors when detecting non-wear or sleep periods. Additionally, the analysis of each file must be performed individually which can be really time-consuming if there is a high number of subjects to be evaluated.

There are two packages in the statistical coding language R (16) that deal with accelerometry data. GENEAREad (17) is designed to load data from GENEActiv devices.while GGIR (18) also processes, calibrates and extracts results from it. This package allows parameter customization but requires a skilled operator to generate and interpret the results.

These problems highlight the relevance of achieving coding and statistical literacy. The multidisciplinary characteristic of this problem requires an efficient communication between

health or physical activity professionals and computer scientists or statisticians. A mutual understanding allows a holistic approach that will align the software calculations with the research needs.

In addition to the mentioned concerns with current software for accelerometry data evaluation, the optimization of the detection of non-wear intervals. A misinterpretation of non-wear time could either underestimate or overestimate sedentary time and therefore lead to an erroneous measure of the actual physical activity participation.

A general overestimation of non-wear time occurs when using the software provided by GENEActiv, with reported non-wear intervals of up to eight hours per day in subjects who actually wore the accelerometer during the whole register. Another method to detect these non-wear periods, proposed by van Hees et al. (19), is based on the analysis of the standard deviations and ranges of each of the three axis during 30-minute intervals. The non-wear cut-off point is set at 30 mg for standard deviations and 50 mg for ranges. However, these thresholds may not be optimal for every population or sampling frequencies different from those used by the authors (40 and 80 Hz).

Therefore, the main purpose of this study was to develop an original software to process and analyze raw accelerometry data in a quick and versatile way. An additional goal was to examine the precision of the thresholds for the detection of non-wear intervals.

2. Methods

The software was developed in R, a coding language specifically designed for statistical computing (16). The use of RStudio, an integrated development environment greatly simplified code development, testing and handling. Coding literacy was achieved through completion of several online courses and an extensive tutorial built within the R package *swirl* (20).

To extend the base functionalities of R, a series of packages were included in the development of the code. More specifically, *rio* (21) was used to manage data input and output, *dplyr* (22) and *data.table* (23) to optimize data formatting and storage, *lubridate* (24) to format dates and times, *magrittr* (25) to concatenate functions, *rmarkdown* (26) to generate reports, *ggplot2* (27) to create plots, *miniGUI* (28) to add a graphic user interface (GUI) and *roxygen2* (29) to create the code documentation.

The decision of the functions that the program should include and the parameters that the researchers should be able to modify was based on the functionalities of the previously existing software, GENEActiv software and *GGIR* package (18), as well as with the collaboration of researchers that are currently carrying out accelerometry measurements.

In order to test the software, 35 participants wore GENEActiv accelerometers on their non-dominant wrist during a week. Eight of the accelerometers were configured to record at 30 Hz, while the sampling frequency was reduced to 10 Hz for the rest of them.

After noticing an underestimation of the non-wear time using the method and cutpoints proposed by van Hees et al. (19), a statistical evaluation of the wear and non-wear intervals of one participant was carried out.

Following the same approach as the aforementioned authors, standard deviations and ranges for each axis during 30-minute intervals were calculated. These intervals were manually labelled as wear or non-wear periods according to the participant annotations. Two different strategies were created to propose new cutpoints.

The first option involved minimizing the type I error by calculating the probability density functions for the standard deviation and range of each axis and setting the threshold at the 95th percentile of the distribution. To optimize both sensitivity and specificity, a second set of cutpoints was created after evaluating the receiver operator characteristics (ROC) curves with the *pROC* package (30).

The reliability of these new cut-off points was checked with the accelerometry registers of other participants.

3. Results

3.1. Software

The final version of the software consists of a total of 765 lines of uncommented code, distributed into 21 functions and a markdown template file. The complete list of functions can be checked in table 1 and the code written in full is included in the digital version of this document. During the development, 32 other functions (537 lines of code) were created but not included in the final version, since they served for testing purposes or were outperformed by the functions finally included. Figure 1 shows the overall structure of the data flow during the analysis process.

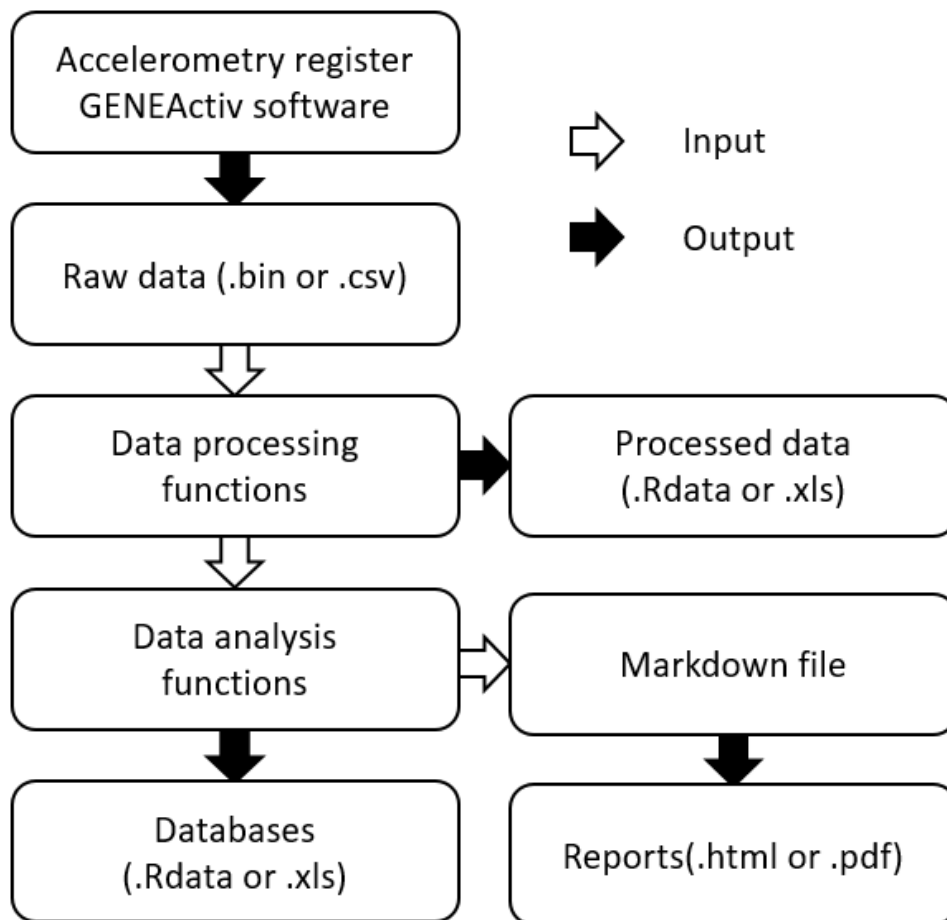
Table 1 – Functions included in the final version of the software

PROCESSING FUNCTIONS	ANALYSIS FUNCTIONS	DIRECTORY FUNCTIONS
<i>Input/Output</i> readCsv (7) readBin (7) writeRData (3)	<i>Activity intensities</i> actZones (7)	<i>Database creation</i> zonedb (85) boutdb (46) sedBoutdb (46)
<i>Data formatting</i> formatDate (5) changeDot (4) setStart (6) setDuration (4)	<i>Bout analysis</i> boutAnalysis (65) sedBoutAnalysis (64)	<i>Report generation</i> completereport (141)*
<i>Non-wear evaluation</i> nonWearEvaluation (30) seeNonWearEvaluation (25)	<i>Most and least active zones</i> M5L5 (36)	<i>Global analysis</i> analysis (49) complete (64) completeMiniGUI (37)
<i>Epoch conversion</i> convertEpoch (20)	<i>Plotting results</i> plotacel (14)	

Numbers in parentheses indicate the lines of uncommented code

*completereport is a markdown file, not a function

Figure 1 – Flow-chart of the analysis



3.1.1. Data processing functions

This group of functions serves to the purpose of loading the data from its original form and prepare it to be later examined. The functions *readCsv* and *readBin* load into R .csv and .bin data respectively, whereas *writeRData* allows to save any result obtained. There are two functions, *formatDate* and *changeDot*, which transform the date variable into the appropriate format for its posterior analysis. The beginning and duration of the register interval can be fixed thanks to the functions *setStart* and *setDuration*.

Two functions carry out the detection of non-wear periods: *nonWearEvaluation* and *seeNonWearEvaluation*. A detailed analysis of this issue will be the topic of a later section.

The data from the three axes is summarized by calculating the Euclidean norm minus one (ENMO). Then, in order to enable a faster analysis, all variables are aggregated into intervals of a fixed duration, denominated “epoch”. These operations are performed with the function *convertEpoch*. The process of development of this function went through various stages as new options to speed up the calculations arose. In table 2 the time needed to group the data into epoch is shown for the different versions of the function.

Table 2. Epoch conversion time (seconds) for different data formats

	1k rows	10k rows	100k rows	1m rows	10m rows
<i>data.table</i>	0	0.02	0.11	0.92	5.40
<i>data.table v.2</i>	0	0.02	0.05	0.73	7.14
<i>data.frame</i>	0.13	2.51	326.8	-	-
<i>data.frame v.2</i>	0.13	2.00	140.05	-	-
<i>dplyr</i>	0.01	0.02	0.08	0.70	7.03

Times are subject to variation from the processing power of the hardware used in the analysis

3.1.2. Data analysis functions

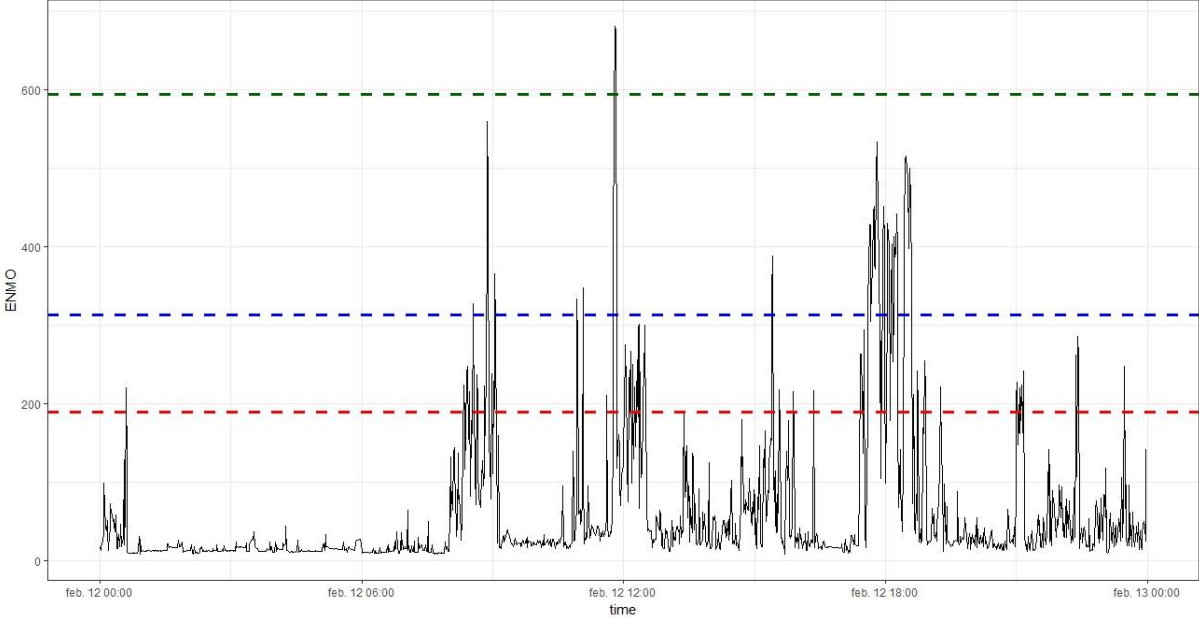
Once the data is properly formatted, a series of functions can be called to obtain results from the accelerometry register. After setting ENMO based cutpoints for light, moderate, and vigorous physical activity, the function *actZones* classifies each epoch according to these categories, which can be later summarized to obtain the amount of time spent in each activity intensity.

To search for consecutive of periods (usually denominated bouts) at a certain activity level, the functions *boutAnalysis* and *sedBoutAnalysis* allow to detect these bouts respectively above or below a previously set threshold.

The function *M5L5* scans the data looking for both the most and the least active periods of time across the register. The duration of the intervals as well as the scan resolution can be modified with the function arguments.

Finally, the function *plotacel* can be called to show a graphical evolution of the ENMO variable across time with visual indicators of the physical activity intensity thresholds. An example of the plot for a single day register is provided in figure 2.

Figure 2 – Graphical representation of the accelerations during a one-day register



Green, blue and red dotted lines represent vigorous, moderate and light physical activity thresholds respectively

3.1.3. Batch file processing

The functions *analysis* and *complete* are designed to perform all previous activities at once. Moreover, a single call to these functions is enough to consecutively analyze all the files in the same directory. These functions manage the folder structure automatically in order to keep the original, processed and analyzed data adequately stored, as it is shown in figure 3. Additionally, the use of the base function *tryCatch* allows to skip a file if the program encounters an error and proceed with the analysis of the next file. The time needed to load and process the data as well as to generate the results, databases and reports for different sampling frequencies can be checked in table 3.

Table 3 – Analysis timing for several files with different sampling frequencies

Sampling frequency (Hz)	Number of files	Epoch duration (s)	Loading time (s)	Processing time (s)	Reporting time (s)	Total time per file (s)
30	8	60	458	334.72	17.72	106.14
10	8	60	146	66.63	16.92	35.87

Based on one-week registers

Figure 3 – Directory structure for data storage



3.1.4. Individual report generation

As part of the consecutive analysis previously described, an individual report is generated for each file analyzed. A customizable markdown template file, named *completereport* serves as a template for the result presentation and automatically includes individual information. This report includes daily results of the time spent in each activity intensity, descriptive statistics of the activity bouts, detection of the most and least active times and a plot featuring the whole register.

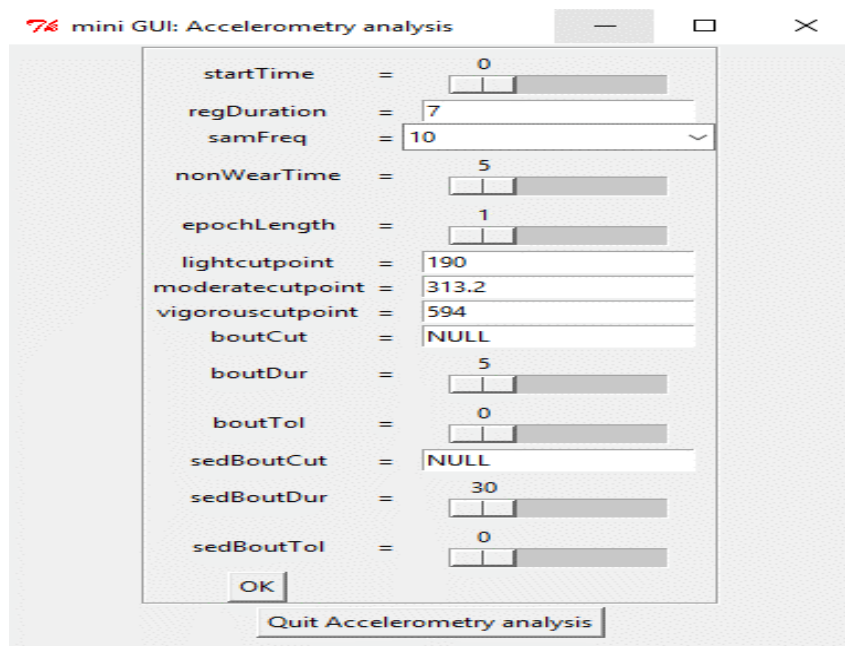
3.1.5. Database creation

In addition to the individual report generation, the *complete* function calls another three functions (*zonedb*, *boutdb* and *sedBoutdb*) that create databases for time spent in each activity intensity, activity bouts and sedentary bouts. Data within these databases is stored as one column for each variable and one row for each included subject. These databases include a set of variables by default, but, if needed new information can be added to match the research requirements.

3.1.6. Graphic user interface

The function *completeMiniGUI* performs the same tasks as the *complete* function but allows an easy introduction of arguments for the functions via a window prompt. The parameters that the user can modify directly from the GUI can be checked in figure 4.

Figure 4 – Window prompt from the graphical user interface



3.2. Non-wear detection

Two sets of cutpoints for the detection of non-wear time, based on the standard deviations and ranges of the acceleration across 30-minute intervals, were created and are shown in table 4. The first set was defined as the 95th percentile of the empirical distribution of such statistics during non-wear time, which are depicted in figure 5, along with the wear time distributions. The second set of cutpoints was based on ROC curves analysis. Both the X-axis and the Y-axis standard deviations showed areas under the curve (AUC) of 1 and allowed a perfect discrimination between wear and non-wear time. The worst classification potential was found in the Y-axis range, showing an AUC of 0.994 which is nonetheless considered as very high (figure 6).

The agreement between the classifications using the new sets of cut-off points and the actual wear status is shown in table 5.

Table 4 – Thresholds for non-wear time detection according to different approaches

	Standard deviation (mg)			Range (mg)		
	X-axis	Y-axis	Z-axis	X-axis	Y-axis	Z-axis
van Hees et al. (19)	30	30	30	50	50	50
p95	6.79	6.98	8.37	105.8	156.1	98.3
ROC	11.6	14.0	11.0	233.5	498.3	393.2

Figure 5 – Standard deviation and range density distributions for wear and non-wear periods

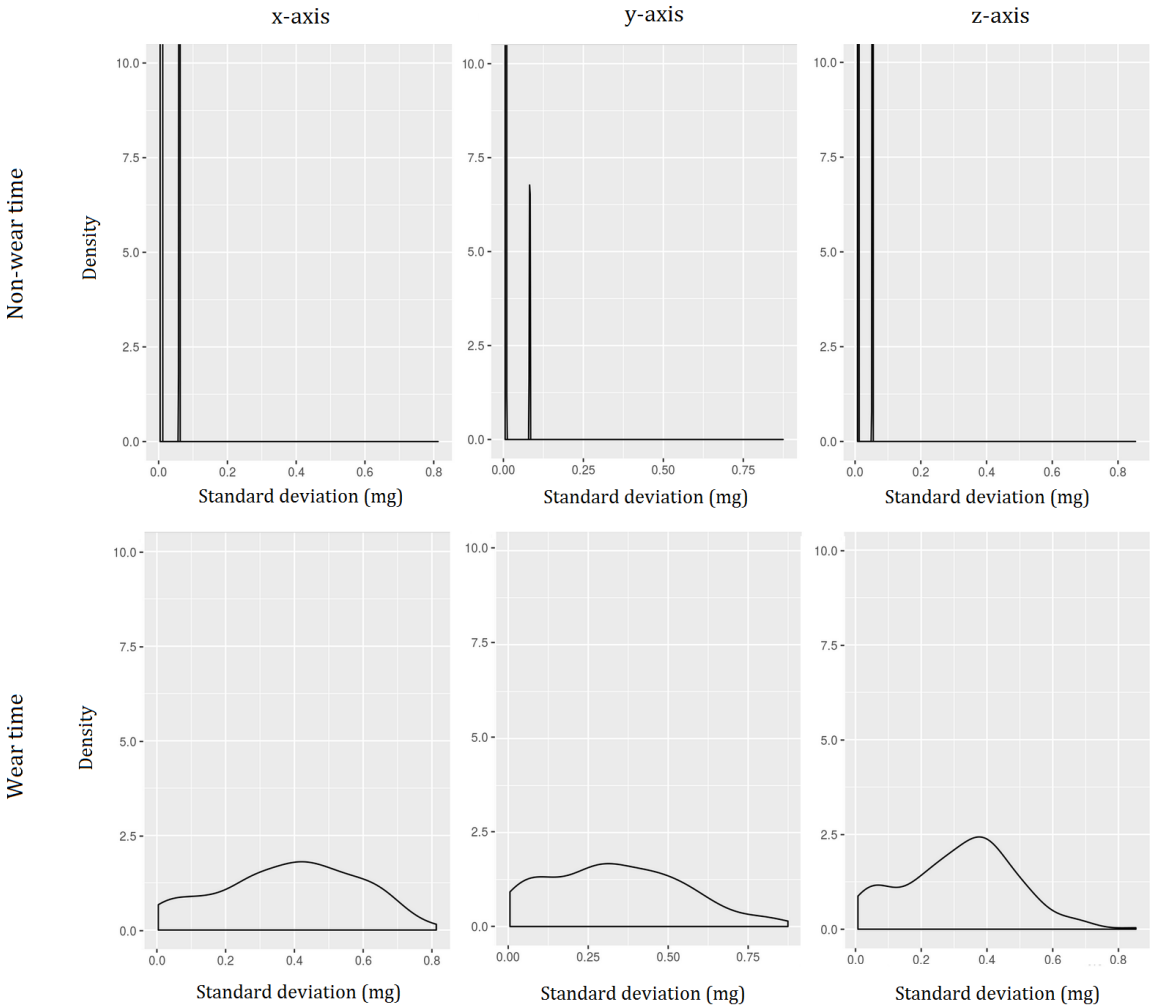
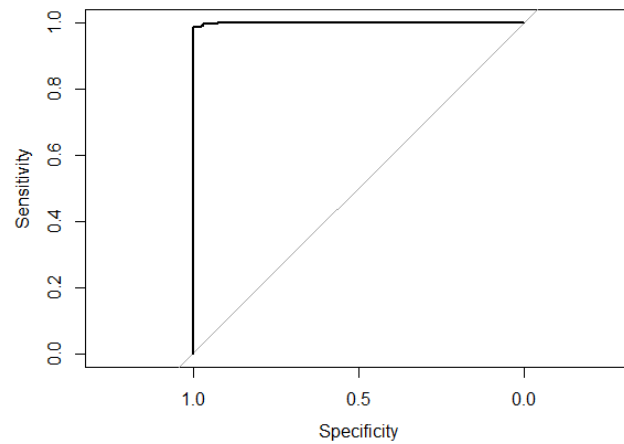


Figure 6 – ROC curve for the determination of non-wear intervals from Y-axis range evaluation



Area under the curve (AUC): 0.994

Table 5 – Non-wear intervals agreement between the different classifications and participant annotations

<i>Training data</i>				
	Wear time (202 intervals)		Non-wear time (136 intervals)	
	Classified correctly	Misclassified	Classified correctly	Misclassified
van Hees et al. (19)	202	0	21	115
p95 cutpoints	201	1	134	2
ROC cutpoints	200	2	136	2
<i>Testing data</i>				
	Wear time (37 intervals)		Non-wear time (83 intervals)	
	Classified correctly	Misclassified	Classified correctly	Misclassified
van Hees et al.	37	0	35	48
p95 cutpoints	37	0	82	1
ROC cutpoints	37	0	83	0

4. Discussion

4.1. Software development

4.1.1. Main features

The newly developed software includes a series of features that might prove useful to researchers interested in physical activity measurement through accelerometry. Its main purpose is the conversion of raw accelerometry data into useful information for both researchers and users.

The primary output from the software is a series of structured databases including both activity intensity and bout duration outcomes that are immediately ready for statistical analysis. It also generates individual reports for each subject, aimed to provide the participants feedback from the study as well as a briefing for the health sciences researcher.

The main limitations of the currently available software for accelerometry data analysis are either the lack of modifiable parameters or the difficulty for the analysis customization. In order to overcome these problems, the new tool includes options to personalize the process in a user-friendly display. 14 different variables are open to customization in a GUI, the data localization can be manually specified via a window prompt and the folder structure and data storage is managed automatically. These features should help researchers conducting the accelerometry evaluation with the parameters that fit best their research needs, while minimizing the requirement of having programming experience.

As it has been shown in the results, this tool is able to completely analyze a one-week accelerometry register with a sampling frequency of 30 Hz (totaling 18.144.000 rows of raw data) in under two minutes. The use of batch analysis enables to evaluate several files at once. Even though, for a high number of files, this task may take several hours, it does not require attention from the user and, thanks to the inclusion of an error recovery mechanism, the execution is not stopped in the event of an error in a single file, advancing to the next one. This feature allows leaving the program running unattended, and obtaining the results from several registers at once.

Another characteristic worth highlighting is the open-source nature of the platform in which the software has been developed. This tool is completely free to use, thus offering an accelerometry analysis solution to all the research community, regardless of the economic status and thus, helping to expand scientific knowledge. Additionally, the source code can be examined and modified, enabling advanced users to further personalize the analysis, or even add new features to fit their research needs.

4.1.2. Applications

The presented software could help researchers from different areas. A precise measurement of physical activity could classify subjects into activity categories or different populations may be compared in terms of physical activity participation. These classifications or comparisons could be taken into account to determine the effect of physical activity in various conditions.

Additionally, the accelerometry data analysis can be used to determine whether or not a participant complies with the physical activity guidelines which can help detecting subjects with an increased risk of suffering diverse pathologies (4).

It has been shown that a sedentary lifestyle can exert a negative impact on health independently of the practice of physical activity (31,32). The analysis of sedentary bouts included in this software could be useful to assess sedentary patterns and analyze its influence on health status.

4.2. Non-wear detection

A new threshold for the detection of non-wear intervals has been included in the software calculations. The program provided by GENEActiv greatly overestimates in all of the 35 subjects evaluated. On the other hand, the use of the cut-off points proposed by van Hees et al. results in an underestimation of the non-wear time when compared with the participant annotations. This could be due to the fact that the thresholds were determined using sample frequencies of 40 and 80 Hz, whereas our data was recorded at 10 or 30 Hz.

The method used in the proposal of new cut-off points for non-wear time detection is entirely device data driven, by means of a simple measuring experiment. Two different approaches were followed. The use of the 95th percentile of the non-wear time distribution is aimed to control the type I error and requires only non-wear measurements, whereas the ROC curve analysis is designed to optimize specificity and sensitivity together and requires both wear and non-wear measurements.

Both the thresholds based on the frequency distribution function and those derived from ROC curves analysis show a strong agreement with the non-wear periods manually classified. This could be crucial for the accurate evaluation of physical activity and especially for the analysis of sedentary patterns, since a misclassification of non-wear intervals will affect the results of the time spent in sedentary behaviors.

5. Strengths and limitations of the study

The most important features of the software that may help researchers and health professionals are the number and diversity of the variables that can be customized in the analysis in a simple way and its management of multiple files at once, generating both individual reports and grouped databases. Also, the open-source nature of the software will allow researchers to use this tool for free and to further personalize the analysis.

Additionally, the proposal of new methods and cut-off points for detecting non-wear intervals can help determining precisely the time periods in which the data does not reflect the actual physical activity of the subject.

One of the main limitations of this study is the lack of an autocalibration algorithm for the raw data, which would help ensuring an optimal sensor precision. Moreover, the thresholds for the

non-wear intervals were obtained from a small sample size and should be validated with a broader variety of subjects and conditions.

The validation of the proposed cutpoints represents an opportunity for future research as well as the inclusion of new functions to examine different activity patterns such as sleep time or sedentarism.

6. Conclusions

A new tool for accelerometry data analysis has been developed. It will enable researchers to easily and quickly obtain results from accelerometer registers. The software includes a user-friendly personalization of the analysis to fit different research methodologies and automatically generates databases and individual reports from multiple files at once. Furthermore, a new set of cut-off points for non-wear time detection has been established. These new cut-off points allow a precise classification of wear and non-wear intervals.

7. References

1. Warburton DER, Nicol CW, Bredin SSD. Health benefits of physical activity: the evidence. *Can Med Assoc J.* 2006;174(6):801–9.
2. Pedersen BK, Saltin B. Exercise as medicine - Evidence for prescribing exercise as therapy in 26 different chronic diseases. *Scand J Med Sci Sport.* 2015;25:1–72.
3. U.S. Department of Health and Human Services. Physical activity guidelines for Americans. *Pres Counc Phys Fit Sport Res Dig [Internet].* 2008;9(4):1–8. Available from:
www.health.gov/paguidelines<http://www.health.gov/paguidelines/guidelines/#toc>
<http://www.health.gov/paguidelines/pdf/paguide.pdf><http://www.health.gov/paguidelines>
4. World Health Organization. Global recommendations on Physical Activity for health. World Health Organization; 2010.
5. Livesey G, Elia M. Estimation of energy expenditure, net carbohydrate utilization, and net fat oxidation and synthesis by indirect calorimetry: evaluation of errors with special reference to the detailed composition of fuels. *Am J Clin Nutr.* 1988;47(4):608–28.
6. Prentice AM. The doubly-labelled water method for measuring energy expenditure. Technical recommendations for use in humans. International Atomic Energy Agency; 1990.

7. Craig CL, Marshall AL, Sjöström M, Bauman AE, Booth ML, Ainsworth BE, et al. International physical activity questionnaire: 12-country reliability and validity. *Med Sci Sport Exerc.* 2003;35(8):1381–95.
8. Folsom AR, Jacobs DR, Caspersen CJ, Gomez-Marin O, Knudsen J. Test-retest reliability of the Minnesota leisure time physical activity questionnaire. *J Chronic Dis.* 1986;39(7):505–11.
9. Hagströmer M, Bergman P, De Bourdeaudhuij I, Ortega FB, Ruiz JR, Manios Y, et al. Concurrent validity of a modified version of the International Physical Activity Questionnaire (IPAQ-A) in European adolescents: The HELENA Study. *Int J Obes.* 2008;32:S42–8.
10. Jacobs Jr DR, Ainsworth BE, Hartman TJ, Leon AS. A simultaneous evaluation of 10 commonly used physical activity questionnaires. *Med Sci Sports Exerc.* 1993;25(1):81–91.
11. Janz KF. Validation of the CSA accelerometer for assessing children’s physical activity. *Med Sci Sport Exerc.* 1994;
12. Sasaki JE, John D, Freedson PS. Validation and comparison of ActiGraph activity monitors. *J Sci Med Sport.* 2011;14(5):411–6.
13. Evenson KR, Catellier DJ, Gill K, Ondrak KS, McMurray RG. Calibration of two objective measures of physical activity for children. *J Sports Sci.* 2008;26(14):1557–65.
14. Hänggi JM, Phillips LRS, Rowlands A V. Validation of the GT3X ActiGraph in children and comparison with the GT1M ActiGraph. *J Sci Med Sport.* 2013;16(1):40–4.
15. Esliger DW, Rowlands A V., Hurst TL, Catt M, Murray P, Eston RG. Validation of the GENE accelerometer. *Med Sci Sports Exerc.* 2011;43(6):1085–93.
16. R Core Team. R: A Language and Environment for Statistical Computing [Internet]. Vienna, Austria: R Foundation for Statistical Computing; 2016. Available from: <https://www.r-project.org>
17. Fang Z. GENEaread: Package For Reading Binary files [Internet]. 2013. Available from: <https://cran.r-project.org/package=GENEaread>
18. Van Hees VT. GGIR: Raw Accelerometer Data Analysis [Internet]. 2017. Available from: <https://cran.r-project.org/package=GGIR>

19. van Hees VT, Renström F, Wright A, Gradmark A, Catt M, Kong YC, et al. Estimation of Daily Energy Expenditure in Pregnant and Non-Pregnant Women Using a Wrist-Worn Tri-Axial Accelerometer. *PLoS One*. 2011;6(7).
20. Kross S, Carchedi N, Bauer B, Grdina G. swirl: Learn R, in R [Internet]. 2016. Available from: <https://cran.r-project.org/package=swirl>
21. Chan C, Chan GC, Leeper TJ, Becker J. rio: A Swiss-army knife for data file I/O. 2016.
22. Wickham H, Francois R. dplyr: A Grammar of Data Manipulation [Internet]. 2016. Available from: <https://cran.r-project.org/package=dplyr>
23. Dowle M, Srinivasan A. data.table: Extension of “data.frame” [Internet]. 2016. Available from: <https://cran.r-project.org/package=data.table>
24. Grolemund G, Wick. Dates and Times Made Easy with lubridate. *J Stat Softw* [Internet]. 2011;40(3):1–25. Available from: <http://www.jstatsoft.org/v40/i03/>
25. Milton Bache S, Wickham H. magrittr: A Forward-Pipe Operator for R [Internet]. 2014. Available from: <https://cran.r-project.org/package=magrittr>
26. Allaire J, Cheng J, Xie Y, McPherson J, Chang W, Allen J, et al. rmarkdown: Dynamic Documents for R [Internet]. 2016. Available from: <https://cran.r-project.org/package=rmarkdown>
27. Wickham H. ggplot2: Elegan Graphics for Data Analysis [Internet]. New York: Springer-Verlag; 2009. Available from: <http://ggplot2.org>
28. Ojeda-Cabrera JL. miniGUI: tktcl quick and simple function GUI. 2010.
29. Wickham H, Danenberg P, Eugster M. roxygen2: In-Line Documentation for R [Internet]. 2017. Available from: <https://cran.r-project.org/package=roxygen2>
30. Robin X, Turck N, Hainard A, Tiberti N, Lisacek F, Sanchez J-C, et al. pROC: an open-source package for R and S+ to analyze and compare ROC curves. *BMC Bioinformatics*. 2011;12:77.
31. Ivuskans A, Maestu J, Jurimae T, Latt E, Purge P, Saar M, et al. Sedentary time has a negative influence on bone mineral parameters in peripubertal boys: a 1-year prospective study. *J Bone Miner Metab*. 2015 Jan;33(1):85–92.
32. Shields M, Tremblay MS. Sedentary behaviour and obesity. *Heal Reports*. 2008;19(2):19.