Heliyon 11 (2025) e42154

Contents lists available at ScienceDirect

Heliyon

journal homepage: www.cell.com/heliyon

Research article

Open-source code for analysis of the lifetime total physical activity questionnaire (LTPAQ)

Coral Sánchez-Martín [a,b,c], Javier Leal-Martín [a,b,c], Timo Rantalainen [d], Irene Rodríguez-Gómez [a,b,c,e], Luis M. Alegre [a,b,c], Ignacio Ara [a,b,c,*]

[a] GENUD Toledo Research Group, Faculty of Sport Sciences, University of Castilla-La Mancha, Toledo, Spain
[b] Centro de Investigación Biomédica en Red Fragilidad y Envejecimiento Saludable (CIBERFES), Instituto de Salud Carlos III, Madrid, Spain
[c] Grupo Mixto de Fragilidad y Envejecimiento Exitoso UCLM-SESCAM, Universidad de Castilla-La Mancha-Servicio de Salud de Castilla-La Mancha, IDISCAM, Toledo, Spain
[d] Gerontology Research Center, Faculty of Sport and Health Sciences, University of Jyväskylä, Finland
[e] IGOID Research Group, Faculty of Sport Sciences, University of Castilla-La Mancha, Toledo, Spain

ARTICLE INFO

Keywords:
Lifespan
Energy expenditure
Assessment
MATLAB script
GitHub

ABSTRACT

Advancing research to understand the impact of lifetime physical activity patterns on health is essential. While some progress on the topic has been made, further research is nevertheless warranted. Openly accessible operational resources for processing and analysing longitudinal physical activity data would help advance the field. The Lifetime Total Physical Activity Questionnaire (LTPAQ) provides an in-depth and detailed record of occupational/volunteer, household and recreational physical activities conducted across the lifespan. Despite the authors describing a specific code for physical activity data processing, its lack of comprehensiveness hinders its implementation by researchers or health professionals with limited programming or software expertise. For this purpose, we aimed to develop an open-source MATLAB code to simplify data handling and enhance comparability between studies using the LTPAQ. The source code of the LTPAQ analysis is accessible on our Git repository (https://github.com/coralsanchezmartin/LTPAQ_MATLAB_code). This code automates data handling, ensuring efficiency and facilitating adaptability. Moreover, it interacts seamlessly with Microsoft Excel sheets, providing a user-friendly interface for data input, storage, and organisation. This approach, combined with ongoing efforts to refine the code, seeks to overcome LTPAQ's scoring complexities, and promote a standardised methodology. By facilitating the processing of physical activity data from this questionnaire, our tool is expected to encourage its application in epidemiological studies and foster new insights into the impact of lifestyle from a life course perspective, ultimately playing a crucial role in shaping public health policies and enhancing clinical practice.

1. Introduction

Despite physical activity (PA) being one of the most recognised health-related modifiable lifestyle factors [1], 23 % of adults fail to meet the recommendations established by the World Health Organization [2]. Accurately estimating lifetime exposure to PA is crucial for evaluating associations between PA and disease outcomes, as chronic diseases often result from sustained lifestyle patterns

* Corresponding author. GENUD Toledo Research Group, Universidad de Castilla-La Mancha, Avda. Carlos III s/n, 45071, Toledo, Spain.
E-mail address: ignacio.ara@uclm.es (I. Ara).

https://doi.org/10.1016/j.heliyon.2025.e42154
Received 14 January 2025; Accepted 20 January 2025
Available online 21 January 2025
2405-8440/

established over extended periods, including childhood [3]. However, measuring lifelong PA is challenging. Accelerometers are currently the preferred method for quantifying PA and sedentary behaviour due to their accuracy in measuring frequency, duration, and intensity [4]. Nevertheless, they are limited in assessing lifetime PA exposure due to constraints in storage capacity and battery life, as well as difficulties in ensuring participant adherence and consent over extended periods. In contrast, PA recall questionnaires, such as the Lifetime Total Physical Activity Questionnaire (LTPAQ), are suitable for large-scale population studies due to their cost-effectiveness, ease of administration, and non-invasiveness [5]. The LTPAQ is designed to estimate relatively regular PA performed over the lifespan by capturing respondents' recollection of the frequency, duration, and intensity of various activities across different domains (occupational, household, recreational, and commuting) [6]. This questionnaire expresses PA in terms of metabolic equivalents (METs), that are absolute physical intensity units representing multiples of a standardised resting metabolic rate (mL $O_2 \cdot kg^{-1} \cdot min^{-1}$ or $kcal \cdot h^{-1}$) [6,7]. However, scoring this questionnaire poses significant challenges due to the extensive information it reports. For instance, the SAS code provided in the LTPAQ User Guide from the Department of Cancer Epidemiology and Prevention Research Alberta is illustrative but comes with a disclaimer about its comprehensiveness and applicability [8]. This may help explain the diversity in analytical approaches and methodologies in existing studies, leading to inconsistencies and difficulties in synthesising

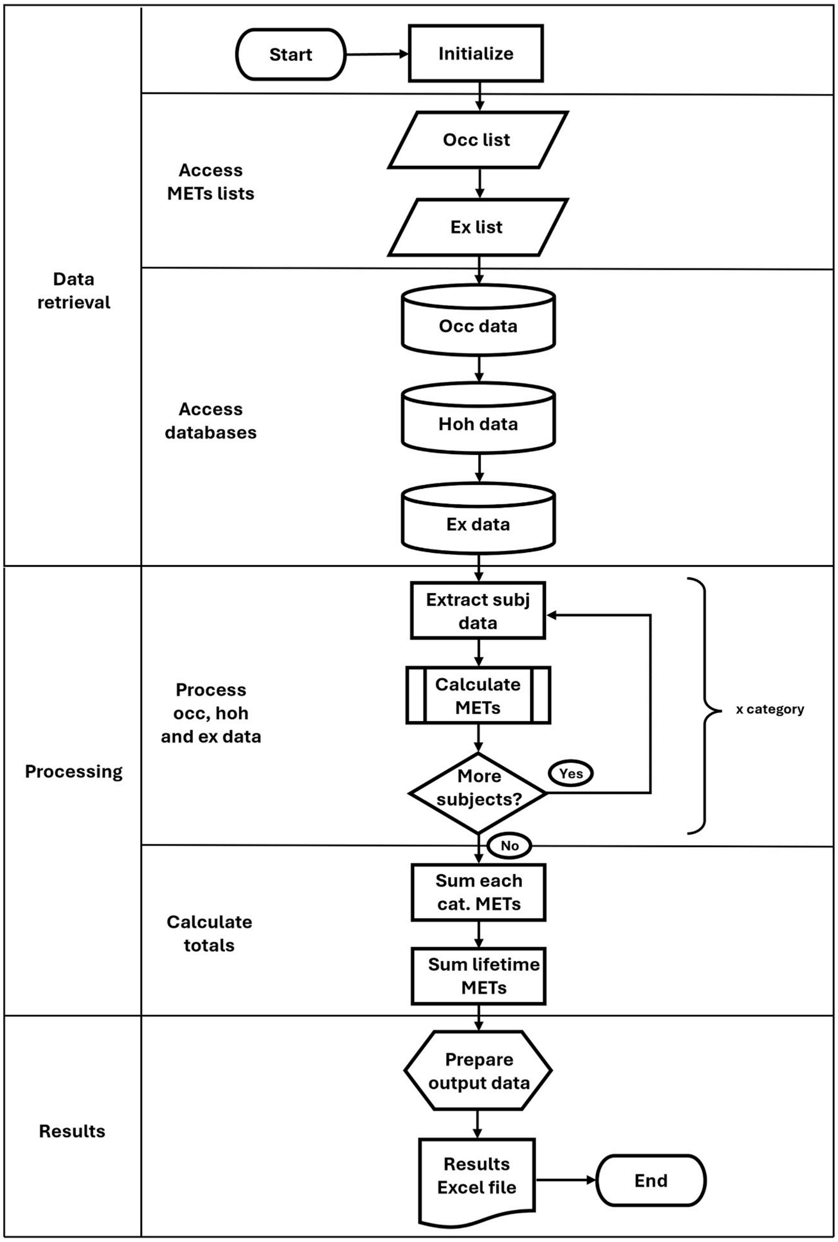

Fig. 1. Flowchart of the data handling steps implemented in the MATLAB code.

data. Some studies have used modified versions of the LTPAQ, reporting only certain activities, focusing on specific target ages, or standardising all household activities at a fixed METs value without differentiating by intensity [9–11]. Others have adhered more closely to the validated LTPAQ methodology but lacked detailed information on data handling and analysis [12]. These limitations emphasise the need for a more standardised and accessible approach to scoring the LTPAQ. Therefore, we aimed to develop an open-source MATLAB code designed to enhance usability, standardise approaches for comparability across studies, and simplify data handling. This effort is crucial for advancing research, promoting a unified understanding of lifetime PA patterns assessed by this questionnaire and facilitating broader adoption and application of this valuable instrument in epidemiological and public health research.

2. Analytic methods

2.1. LTPAQ structure

The questionnaire is divided into three sections dedicated to gathering information on "occupational and volunteer activities", "household activities", and "exercise and sports activities". In all sections, the total duration (in years), frequency (months per year, days per week, and time per day in hours and minutes), and the individuals' perceived intensity for each activity entry are recorded. In the "occupational and volunteer activities" section, alongside the previously mentioned details, it is necessary to encode up to three activities describing the participant's job title. Furthermore, within this section, the questionnaire also captures information on active commuting (means of active transportation and its frequency of use). Regardless of the section, the perceived intensity for each activity entry is categorised by assigning a score of 2 to minimal effort, 3 to activities that are not strenuous but slightly increase heart rate, and 4 to activities that lead to increased heart rate and cause intense sweating. Exceptionally, in the "occupational and volunteer activities" section, a score of 1 is used for jobs that require only sitting with minimal walking.

2.2. LTPAQ data handling

The assessment of PA in this questionnaire involves converting the time spent in each activity entry into METs [8]. The hours dedicated of each activity are multiplied by the assigned METs-based intensity values [13] from the Compendiums of Physical Activities [14–17]. Total PA, in terms of METs$\cdot$h^{-1}, is estimated as the sum of occupational, commuting, household, and exercise/sports activities performed since childhood. This conversion process is automated within the developed MATLAB code. A detailed breakdown of the steps executed by the code is provided in Fig. 1. To facilitate data input, three Microsoft (MS) Excel templates are provided, and detailed instructions on their usage are included as Supplementary Material. These templates are accessible by MATLAB for retrieval and processing. The first spreadsheet serves as the database for questionnaire data, requiring meticulous input to maintain consistency. Each section of the questionnaire is organised on a separate MS Excel tab to improve clarity. Two additional MS Excel files store METs values for various activities, with one sheet dedicated to occupations and another to exercise/sports. METs values for transportation and household activities are included as constants within the code. For household activities, the LTPAQ assigns three predefined METs values corresponding to intensity categories 2 (2.5 METs), 3 (3.5 METs), and 4 (4.5 METs) [8]. In contrast, since the questionnaire lacks specific METs values for the different modes of transportation (walking, biking, rollerblading, and running), we propose the following METs values, based on the Compendium developed by Herrmann, Willis (14): 4, 6.8, 7, and 6 METs, respectively. When the option "Other transportation" is selected, the database spreadsheet includes a cell where the corresponding METs value must be manually entered.

3. Usage considerations

Several factors should be considered when using this tool. Firstly, it is essential to reliably record the physical intensity perceived by the individual. This is important because the same activity can vary in METs values depending on intensity, and such inaccuracies can lead to errors in calculating the overall METs accumulation. Secondly, while our tool ensures consistency in data handling, researchers should consider potential variations in METs values across ages or when involving specific populations. For activities performed during childhood, considering the higher resting metabolic rates per unit of body mass is paramount [18]. Otherwise, assuming standardised adult METs (VO_2/3.5 ml $O_2\cdot kg^{-1}\cdot min^{-1}$) could misinterpret activity intensity in children [19]. Accurate estimates can be achieved using METs from the Youth Compendium of Physical Activities [16]. Conversely, in late adulthood, the resting metabolic rate is lower (2.7 ml $O_2\cdot kg^{-1}\cdot min^{-1}$) [20], so it is advisable to use population-specific METs values ($METs_{60+}$) from the Older Adult Compendium [15]. Additionally, specific METs for individuals who use wheelchairs are available in the Wheelchair Compendium [17].

4. Strengths and limitations

The proposed open-source code offers a standardised and accessible approach for lifetime PA analysis, enhancing comparability across studies. Utilising MATLAB to process data from MS Excel sheets provides distinct advantages, primarily rooted in enhanced efficiency and optimised data handling. Our custom scripts developed using MATLAB enable efficient management of extensive datasets. The integration with MS Excel allows for direct data import and export, minimising manual errors and ensuring a streamlined workflow. Moreover, the use of MS Excel templates is key to the tool's user-friendliness, as it allows researchers to work in a familiar environment, simplifying data input and organisation. This design approach ensures that users can easily navigate the tool, making it accessible to researchers with varying levels of technical expertise. Furthermore, this organised approach provides a high degree of adaptability. METs values can be easily modified in anticipation of Compendium updates without the need to modify the code. This feature ensures alignment with the latest understanding of METs, facilitating a responsive approach to the evolving PA assessment. The code's inherent modifiability allows users to tailor it to their specific needs. For instance, researchers can conveniently adjust how results are expressed or request the retrieval of intermediate variables from the analysis. To ensure accuracy and reliability, we compared the MATLAB code's outcomes with known results and implemented error handling to effectively manage unexpected scenarios. The code is supported with clear documentation and comments that explain calculations and limitations, aiding in future understanding and maintenance. Additionally, it underwent peer review by fellow researchers and domain experts, which helped identify potential issues and areas for improvement.

While the proposed open-source code is a valuable tool for data analysis, it is essential to consider its limitations in application. Firstly, the code requires input data to follow a consistent structure, which is a potential source of errors. However, the provided Excel template helps to minimise these errors by standardising data entry. Secondly, despite the code's effort to enhance standardisation, individual variations in reporting and coding of activities may still pose challenges in achieving complete comparability across studies. Additionally, using predefined METs values for transportation and household activities, may not fully capture individual variations in energy expenditure, despite being aligned with the validated questionnaire. Furthermore, the accessibility and usability of the code may be hindered by researchers' familiarity with MATLAB and its status as a paid software, potentially limiting its widespread adoption. However, leveraging its open-source nature, we encourage future adaptations in other programming languages such as R or Python to enhance accessibility. Finally, to our knowledge, there is no literature available that reports the practical use of the SAS code or describes the detailed data handling of the questionnaire with other available tools, which hinders direct comparisons of our code with existing solutions.

5. Conclusion

In conclusion, this tool assists researchers in processing the LTPAQ by providing a standardised, adaptable, and user-friendly environment, thereby enhancing consistency and comparability across studies. Through this approach to assessing lifetime PA, we aspire to shape a more comprehensive perspective of the lifelong impact of PA. Furthermore, we recognise the importance of ongoing refinement of the code to ensure its effectiveness and broad applicability in the scientific community. Our next steps involve incorporating new variables to offer detailed insights, including age-specific breakdowns and categorisations based on distinct activities or intensities that further enrich the tool's capabilities.

CRediT authorship contribution statement

Coral Sánchez Martín: Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Software, Writing – original draft. **Javier Leal-Martín:** Conceptualization, Data curation, Formal analysis, Methodology, Writing – review & editing. **Timo Rantalainen:** Software, Supervision, Visualization, Writing – review & editing. **Irene Rodríguez-Gómez:** Supervision, Writing – review & editing. **Luis M. Alegre:** Conceptualization, Funding acquisition, Resources, Supervision, Writing – review & editing. **Ignacio Ara:** Funding acquisition, Project administration, Resources, Supervision, Writing – review & editing.

Data repository

We released the source code of the LTPAQ analysis in the Git repository (https://github.com/coralsanchezmartin/LTPAQ_MATLAB_code). The LTPAQ analysis is written in MATLAB and is easily accessible. Detailed instructions for the code are provided in a "readme.md" file available in the Git repository.

Funding

This research was supported by the CIBER de Fragilidad y Envejecimiento Saludable (CIBERFES, group CB16/10/00477), and by the Plan Propio de Investigación of the University of Castilla-La Mancha and FEDER funds from the European Union (2022-GRIN-34296). CSM and JLM received Ph.D. grants from the Universidad de Castilla-La Mancha, Spain (CSM: 2020-PREDUCLM-14936; JLM: 2019-PREDUCLM-11385), and IRG holds a post-doctoral contract from the University of Castilla-La Mancha "Contratos de investigadores postdoctorales para la excelencia científica en el desarrollo del Plan Propio de I + D + i, cofinanciada por el Fondo Social Europeo" (2021/5937). Additionally, CSM received a mobility grant from the Vicerrectorado de Política Científica of the University of Castilla-La Mancha in 2023 (672316. [2023/656]). TR was supported by the Academy of Finland during the preparation of the manuscript (Academy of Finland grant numbers 321336, 328818, and 352653). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Declaration of competing interest

None declared.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.heliyon.2025.e42154.

References

[1] B.K. Pedersen, B. Saltin, Exercise as medicine - evidence for prescribing exercise as therapy in 26 different chronic diseases, Scand. J. Med. Sci. Sports 25 (Suppl 3) (2015) 1–72, https://doi.org/10.1111/sms.12581.

[2] W.H. Organization, Global Action Plan on Physical Activity 2018–2030: More Active People for a Healthier World, World Health Organization, Geneva, 2018.

[3] J. Juhola, M. Oikonen, C.G. Magnussen, V. Mikkilä, N. Siitonen, E. Jokinen, et al., Childhood physical, environmental, and genetic predictors of adult hypertension: the cardiovascular risk in young Finns study, Circulation 126 (4) (2012) 402–409, https://doi.org/10.1161/circulationaha.111.085977.

[4] K.R. Westerterp, Assessment of physical activity: a critical appraisal, Eur. J. Appl. Physiol. 105 (6) (2009) 823–828, https://doi.org/10.1007/s00421-009-1000-2.

[5] C.M. Friedenreich, K.S. Courneya, H.K. Neilson, C.E. Matthews, G. Willis, M. Irwin, et al., Reliability and validity of the past year total physical activity questionnaire, Am. J. Epidemiol. 163 (10) (2006) 959–970, https://doi.org/10.1093/aje/kwj112.

[6] C.M. Friedenreich, K.S. Courneya, H.E. Bryant, The lifetime total physical activity questionnaire: development and reliability, Med. Sci. Sports Exerc. 30 (2) (1998) 266–274, https://doi.org/10.1097/00005768-199802000-00015.

[7] E.T. Howley, Type of activity: resistance, aerobic and leisure versus occupational physical activity, Med. Sci. Sports Exerc. 33 (6 Suppl) (2001) S364–S369, https://doi.org/10.1097/00005768-200106001-00005, discussion S419-20.

[8] Services CeaPRCCAH. Lifetime Total Physical Activity Questionnaire Comprehensive User's Guide 2016.

[9] Y.M. Choe, G.H. Suh, B.C. Lee, I.G. Choi, H.S. Kim, J.W. Kim, et al., High-intensity walking in midlife is associated with improved memory in physically capable older adults, Alzheimer's Res. Ther. 15 (1) (2023) 143, https://doi.org/10.1186/s13195-023-01293-8.

[10] K. Boisvert-Vigneault, H. Payette, M. Audet, P. Gaudreau, M. Bélanger, I.J. Dionne, Relationships between physical activity across lifetime and health outcomes in older adults: results from the NuAge cohort, Prev. Med. 91 (2016) 37–42, https://doi.org/10.1016/j.ypmed.2016.07.018.

[11] J. Kruk, Lifetime physical activity and the risk of breast cancer: a case-control study, Cancer Detect. Prev. 31 (1) (2007) 18–28, https://doi.org/10.1016/j.cdp.2006.12.003.

[12] S.J. Gill, C.M. Friedenreich, T.T. Sajobi, R.S. Longman, L.L. Drogos, M.H. Davenport, et al., Association between lifetime physical activity and cognitive functioning in middle-aged and older community dwelling adults: results from the brain in motion study, J. Int. Neuropsychol. Soc. 21 (10) (2015) 816–830, https://doi.org/10.1017/s1355617715000880.

[13] A.P. Hills, N. Mokhtar, N.M. Byrne, Assessment of physical activity and energy expenditure: an overview of objective measures, Front. Nutr. 1 (2014) 5, https://doi.org/10.3389/fnut.2014.00005.

[14] S.D. Herrmann, E.A. Willis, B.E. Ainsworth, T.V. Barreira, M. Hastert, C.L. Kracht, et al., 2024 Adult Compendium of Physical Activities: a third update of the energy costs of human activities, J Sport Health Sci 13 (1) (2024) 6–12, https://doi.org/10.1016/j.jshs.2023.10.010.

[15] E.A. Willis, S.D. Herrmann, M. Hastert, C.L. Kracht, T.V. Barreira, J.M. Schuna Jr., et al., Older Adult Compendium of Physical Activities: energy costs of human activities in adults aged 60 and older, J Sport Health Sci 13 (1) (2024) 13–17, https://doi.org/10.1016/j.jshs.2023.10.007.

[16] N.F. Butte, K.B. Watson, K. Ridley, I.F. Zakeri, R.G. McMurray, K.A. Pfeiffer, et al., A Youth Compendium of physical activities: activity codes and metabolic intensities, Med. Sci. Sports Exerc. 50 (2) (2018) 246–256, https://doi.org/10.1249/mss.0000000000001430.

[17] S.A. Conger, S.D. Herrmann, E.A. Willis, T.E. Nightingale, J.R. Sherman, B.E. Ainsworth, Wheelchair Compendium of Physical Activities: an update of activity codes and energy expenditure values, J Sport Health Sci 13 (1) (2024) 18–23, https://doi.org/10.1016/j.jshs.2023.11.003, 2024.

[18] Z. Wang, High ratio of resting energy expenditure to body mass in childhood and adolescence: a mechanistic model, Am. J. Hum. Biol. 24 (4) (2012) 460–467, https://doi.org/10.1002/ajhb.22246.

[19] R.G. McMurray, N.F. Butte, S.E. Crouter, S.G. Trost, K.A. Pfeiffer, D.R. Bassett, et al., Exploring metrics to express energy expenditure of physical activity in Youth, PLoS One 10 (6) (2015) e0130869, https://doi.org/10.1371/journal.pone.0130869.

[20] J. Leal-Martín, M. Muñoz-Muñoz, S.K. Keadle, F. Amaro-Gahete, L.M. Alegre, A. Mañas, et al., Resting oxygen uptake value of 1 metabolic equivalent of task in older adults: a systematic review and descriptive analysis, Sports Med. 52 (2) (2022) 331–348, https://doi.org/10.1007/s40279-021-01539-1.