Multivariate Analysis of Adaptation Level in Low-Cost Lower Limb Prostheses: An Unsupervised Learning Approach

Gabriel Maldonado Colmenares(✉) and Jenny Kateryne Nieto Aristizabal(✉)

Bioinstrumentation and Clinical Engineering Research Group – GIBIC, Bioengineering Department, Engineering Faculty, Universidad de Antioquia UdeA, Calle 70 No. 52-21, Medellín, Colombia

{gabriel.maldonado,jenny.aristizabal}@udea.edu.co

Abstract. Objective: To develop an unsupervised learning approach to study the prosthesis adaptation process using functional assessment tests and their scales, health-related behaviors, and socio-demographic da-ta.

Subjects: 199 low-cost lower limb prosthesis users with below-knee and/or above-knee amputation.

Methods: For the unsupervised learning approach, different methods, such as K-Means, Agglomerative Clustering, and Fuzzy C-Means, were used to comprise clusters and classify individuals based on factors associated with the lower limb prosthesis. Davies Boulding, Dunn, and Calinski-Harabasz index as well as the Silhouette coefficient were used to validate, study, and understand the resulting clusters from the dataset.

Results: The unsupervised learning approach strategies provided patient phenotyping clusters that could be interpreted as adaptation levels in low-cost lower limb prosthesis users, while allowing the interpretation and patient phenotyping by physicians.

Conclusions : Patient care customization is important, especially in multidimensional problems. To do so, it is necessary to use historic-data-based tools, which allow better control of the current state of the prosthesis and the person's functional capacity.

Keywords: Adaptation · Amputees · Functional assessment tests · Lower-limb prostheses · Clustering analysis · PEQ-MS · Houghton · 2MWT

1 Introduction

Currently, only 10% of amputees in the world can access a prosthesis, and the population that acquires low-cost technology reports low adaptation and low adherence to the use of these accessible devices [1]. The main cause of this problem is that the diagnosis and use of prosthetic devices require both physiological and psychosocial factors along design parameters, which turn the rehabilitation process into a multidimensional [2]. Besides people's mobility being affected after lower limb amputation, there are several

J. C. Figueroa-García et al. (Eds.): WEA 2021, CCIS 1431, pp. 147–159, 2021.
https://doi.org/10.1007/978-3-030-86702-7_13

side effects, such as body image self-perception as well as lower participation in social and everyday activities: additionally, these factors impact the rehabilitation process [3]. Furthermore, patients from low-income countries tend to use low-cost prosthesis devices: devices costing less than 600 US dollars; moreover, there is a deficient follow-up after the initiation of prosthesis [4, 5]. Another common drawback is that there is no well-defined protocol to assess the quality and level of adaptation in these patients or a defined methodology to predict their walking ability for preventing the abandonment of the prosthesis device [3].

An efficient way to measure the current state of rehabilitation is through functional assessment scales in post-prosthetic rehabilitation, which are direct indicators of the patient's mobility capacity [6]. This tool can estimate an individual's potential to walk with a prosthesis, and nowadays, it is used by different health centers in Colombia to support the monitoring of the patient's post-prosthetic rehabilitation. These assessment scales can be obtained through self-formulated questionnaires such as the Prosthesis Evaluation Questionnaire-Mobility Scale (PEQ-MS) and Houghton questionnaire as well as performance tests such as the Two Minute Walk Time(2MWT) and Time Up and Go (TUG). Each functional assessment test provides valuable information about both the functional capacity and adaptation evaluation of the person undergoing prosthetic rehabilitation [7]. However, due to the high dependency of environmental factors along the population´s characteristics, such as the different amputation cause in the Colombian population, currently, there is no well-defined standard on how to interpret and correlate different functional assessment scales in different settings specifically with low-cost prostheses [6, 7].

This lack of a standardized process makes it difficult to interpret the data collected in clinical practices and correlate it with other influencing factors such as age, physical state, health-related behaviors, and socio-economic level [2]. Given the complexity of simultaneously interpreting many variables and factors, there is a need to create tools that allow for the analysis and interpretation of these data, while extracting valuable information for the specialists. Traditionally, some of the most common analytical methods have been: Descriptive analysis, discriminant analysis, multivariate regression analysis, analysis of covariance (MANCOVA), and analysis of variance (ANOVA). These methods have been used to predict the walking ability after lower limb amputation. However, these approaches are based on a high variety success criterion such as Houghton Questionnaire Score, Daily use of prosthesis, Walking distance, TUG Score, Rivermead Mobility Index and more. In the same way, different predictive factors have been accounted for, including the cause of amputation, the amputation level, stump factors, Body Mass Index (BMI), motivation, occupation, sex, co-morbidities, psychosocial factors, social support, and other health-related behaviors [2]. This heterogeneity of methodologies, population characteristics, inclusion criteria, and outcome measures make the comparison between methods and the management of rehabilitation plans difficult in countries with populations such as Colombia.

When there is a clear need for patient phenotyping, disease subtyping, and adverse events detection; the clustering algorithms acquire great value for the development of new clinical applications [8]. K-Means, Agglomerative Clustering, Gaussian Mixture Models, and Fuzzy C-Means are unsupervised learning methods that have been used for pre-processing and pattern extraction in clinical and health-related data [9]. Besides, these techniques can determine which variables are significantly related to different outcomes like satisfaction, walking ability, and range of motion after amputation [10]. Additionally, there is evidence that unsupervised learning methods are an excellent tool for different telemedicine applications [11, 13]. Nonetheless, there is a lack of studies that applied cluster analysis in problems related with the amputee care and healthcare systems based on questionnaires [14].

This study proposes an unsupervised learning approach based on the prosthetic adaptation level considering different qualities, materials, prices, patient phenotype, among others. through clustering and statistical analysis using the PEQ-MS, Houghton, 2MWT functional assessment scores, as predictive factors.

2 Methodology

2.1 Description of Data

The population of the study was obtained from a cross-sectional study in a sample of 199 persons with lower limb amputation and users of low-cost prosthesis acquired from Mahavir Kmina: a non-profit corporation located in Colombia that helps amputee people regain their mobility and quality of life with free low-cost lower limb prosthesis devices. The data was obtained through a survey conducted to beneficiaries when they attended the corporation for device adjustment, maintenance or in a need of a prosthesis replacement, between the years 2018 and 2020. The survey consists in 42 questions that included basic demographic, medical history information, and functional assessment tests and scales.

Instances were included without age, sex, race, and amputation level restriction. The general inclusion criteria were: Participants that used low-cost lower-limb prostheses to any period and acquire the prosthesis in Mahavir Kmina Corporation made of high-density polyethylene with the foot of Jaipur. Instances were excluded in case of being first-time users without device, presenting sensory-perceptual or cognitive alteration, refusing the informed consent, and instances with very incomplete or inconsistent answers. This informed consent was designed and approved by the technical direction and ethics committees of the Mahavir Kmina Corporation and the Health Rehabilitation group of the University of Antioquia.

The design of the surveys used in the study is contemplated within the clinical practice guidelines established by the Ministry of Health and Social Protection of Colombia [6]. These guidelines suggest the use of the functional assessment tests collected in the survey (PEQ-MS, Houghton Scale, and 2MWT test) as parameters to evaluate specific aspects in the lower limb prosthesis adaptation process like mobility, self-perception and prothesis use. On the other hand, as they are factors that influence the person's daily physical activity, age, sex, amputation cause and level, occupation and the type of supports they use (walking sticks, walkers or not using them) were included the analyzes [3, 7, 15]. The description of the variables collected in the study are disposed in the Table 1.

Table 1. Description and information of the data collected.

Feature name	Description
Age	Age of the prosthesis user
Sex	Sex of the prosthesis user
Independence	3 self-reported levels over 3 different activities of daily living: 1: Independent, 2: Semi-independent and 3: Dependent
Socioeconomic level	Socioeconomic level of the user: 1: Low-low, 2: Low, 3: Medium-low, 4: Medium, 5: Medium-high, 6: High
Daily use time	Average daily prothesis wearing time in hours
Adaptation time	Months since the start of prosthesis adaptation
Satisfaction	General level of satisfaction with the prosthesis extracted from a visual analog scale (VAS): 0 is very unsatisfied and 100 is very satisfied
Occupation	Employment status or principal activity:
Cause of amputation	3 categories were defined: Disease: (e.g., Infections, cancer, diabetes, gangrene), nonviolent accidents (e.g., car or motorbike accident, work accident) and violent accidents (e.g., gunshot, explosive mine)
Amputation level	Amputation level defined in two general categories: Below knee (BK) and above knee (AK)
Assistances user	Type of supports they use: 1: walking sticks, 2: walkers, 3: wheelchair and 4: not using them
PEQ-MS	Score of Prothesis Evaluation Questionnaire- Mobility Scale: 13 self-reported mobility related questions extracted from the original PEQ. The score is extracted from a VAS where: 0 is unable and 100 is fully capable. This test "has demonstrated good psychometric characteristics for measuring mobility and excellent reliability" [16]
Houghton score	Score of Houghton Scale: 4 self-reported questions that is widely suggested and defined as standard in the literature for classify people after the initial prosthetic rehabilitation according to walking ability category Scores $\geq$ 9: Independent community walking ability, scores from 6 to 8: Household and limited community walking ability, and scores $\leq$ 5: Limited-household walking ability [7, 16, 17]
2MWT	Result of the 2 min' walk time functional ability test: Total distance walked by the user in a period of 2 min by requesting their highest walking speed beforehand. This test has excellent test-retest reliability and internal consistency with self-reported functional ability scales and performance assessments like the Houghton score [16]

2.2 Data Analysis

Python and R were used to carry out the statistical analyzes and to explore and preprocess the data. The analysis was developed following the following flow of activities: data cleaning, exploratory and statistical data analysis, unsupervised analysis and cluster's validation, visualization, and interpretation.

Data Cleaning. The database includes multiple instances with null data, typos and inconsistent instances that were remove. The data processing pipeline includes format errors correction, automatic outliers using isolation forest and dimensional reduction analysis, imputation techniques based on regression analysis and over sampling method to maintain the volume of data in the study.

Exploratory and Statistical Data Analysis. The main objective of this stage is to generate conclusions from the data and guide the following data processing steps like feature extraction, outliers detection, and data cleaning and imputation [18]. In this stage, the level of satisfaction of a prosthesis user was used as base line model due to the high influence of satisfaction in the adaptation level [2, 3, 19]. Shapiro Wilks and Kolmogorov tests were used to test the normal distribution in the features and stratified sampling strategies were used from this distribution to generate three categories from the level of satisfaction with his prosthesis expressed by the person. From these proposed levels ANOVA and MANCOVA was used in the following sections to establish the features importance in the overall satisfaction level discrimination using a significance level of p-value = 0.05 to identify statistically significant associations [20], taking into account that the features analyzed were considered as ratio data [16] and previous works have evidenced through univariate techniques such as Chi-square tests, Student's t-test and Spearman's correlation the relevance of the understanding of this associations in low-cost lower limb prosthesis users [21].

Moreover, a correlation analysis was useful to find the associations between features and redundant information in the inner questions of the test. Categorical associations were defined with the uncertainty coefficient & correlation ratio and numerical correlations were defined with the Spearman correlation coefficient. Past studies evidence strong correlations between the 2MWT test and the Houghton scale, and moderate associations between 2MWT test and the PEQ-MS, and evidence a need to propose statistical models that allow a better understanding of the relationship of the tests to each other [7]. The correlation analysis was extended to the individual questions of the PEQ-MS to evaluate associations between specific aspects of the ambulation with the prosthesis and other functional assessment tests. To reduce dimensionality of the dataset, evaluate multiples preprocessing pipelines and facilitate the visualization and interpretation of the results a dimensionality reduction analysis was tested with Principal Components Analysis [22] and u-MAP algorithm [23]. Also, the literature suggest lower dimensionality for better results in the modeling stage [24].

Unsupervised Analysis and Cluster's Validation. Unsupervised learning analysis has been applied for such varied objectives as model fitting, prediction based on groups, hypothesis generation and testing, data exploration, dimension reduction, and grouping similar entities into homogeneous groups [12, 25, 26]. In the case of prosthesis rehabilitation process the literature report a clear group definition and patient phenotyping between young and fitness prosthesis users and the advanced age users, this because there is strong evidence that younger age at amputation results in superior walking ability, however it is still possible for individuals over 90 years of age to walk independently following lower limb amputation [27]. Transition stages or subgroups in the prosthesis adaptation process, and in the clinical practice in general, can represent valuable information to build a custom rehabilitation process according to user needs, and also open the possibility of creating new telemedicine tools for prosthesis users [12, 28]. The main objective of applying these algorithms is find states of interest, transition stages or subgroups in the adaptation process of low-cost lower limb prostheses users, and assets the influence of the different features used in this study over the conformed clusters. Agglomerative hierarchical clustering, K-means, Gaussian mixture models and Fuzzy C-means algorithms were applied in three versions of the dataset*: 1. All patients, 2. Below knee amputee level, and 3. Above knee amputee level* and the following internal validation metrics were used to evaluate the performance of clustering models and find the optimal number of clusters for each case: *1. Davies Boulding Index, 2. Callinsky-Harabaz Index, 3. Silhouette coefficient and 4. Dunn score* [26, 28, 29].

Visualization and Interpretation. The conformed clusters were visualized using the dimension reduction steps constructed in the preprocessing pipeline, PCA transformation and u-MAP embedding were used to visualize the distribution of the clusters and evaluate in a visual way the separability and compaction of these groups [30]. In addition, violin plots were used to study the distribution of specific features within groups and ANOVA tests were carried on to assets the significance level between the features and the conformed groups.

3 Results

Shapiro wilks and Kolmogorov test demonstrated the normal distribution of variables as satisfaction, 2MWT and PEQ-MS with 95% confidence and the Houghton Scale was transformed to a normal distribution with a box cox transformation to fulfill the assumptions of the ANOVA analysis. The Fig. 1 shows the associations map result of the correlation analysis, this representation allows a graphic understanding of the relations between numerical and categorical features used in this study. Squares are categorical associations representing uncertainty coefficient and correlation ratio (0 to 1), circles are the symmetrical numerical associations representing the Spearman correlation values (−1 to 1), and the diagonal is left blank for clarity. Following criterions reported in the literature a weak correlation was defined between absolutes values of 0 and 0.29, moderate between 0.30 and 0.49 and strong between 0.5 and 1 [7].

Satisfaction shows a strong influence by the Houghton Score ($r = 0.51$) and the main occupation/activity of the user (0.52); Moderate influence using assistances (e.g., walking stick, wheelchair) ($r = 0.49$), the PEQ-MS score ($r = 0.40$), daily hours of use ($r = 0.36$) and independence level ($r = 0.30$). The Houghton scale shows a strong positive correlation with daily hours of use ($r = 0.53$), Satisfaction ($r = 51$), and the PEQ-MS score ($r = 0.40$) and moderate correlations with the 2MWT test ($r = 0.37$), age ($r = 0.32$) and amputation cause ($r = 0.31$). In other hand, the PEQ-MS test evidence strong associations with the Houghton score ($r = 0.44$), Satisfaction ($r = 0.40$), and weak associations with de daily hour of use ($r = 0.29$) and the assistances use ($r = 0.21$).

Finally, the 2MWT shows only moderate correlations with Houghton score (r = 0.37) and weak associations with the assistances use (r = 0.29), age (r = −0.26), independence level (r = 0.26), occupation (r = 0.29), socio economic level (r = 0. 29) and satisfaction (r = 0.21).

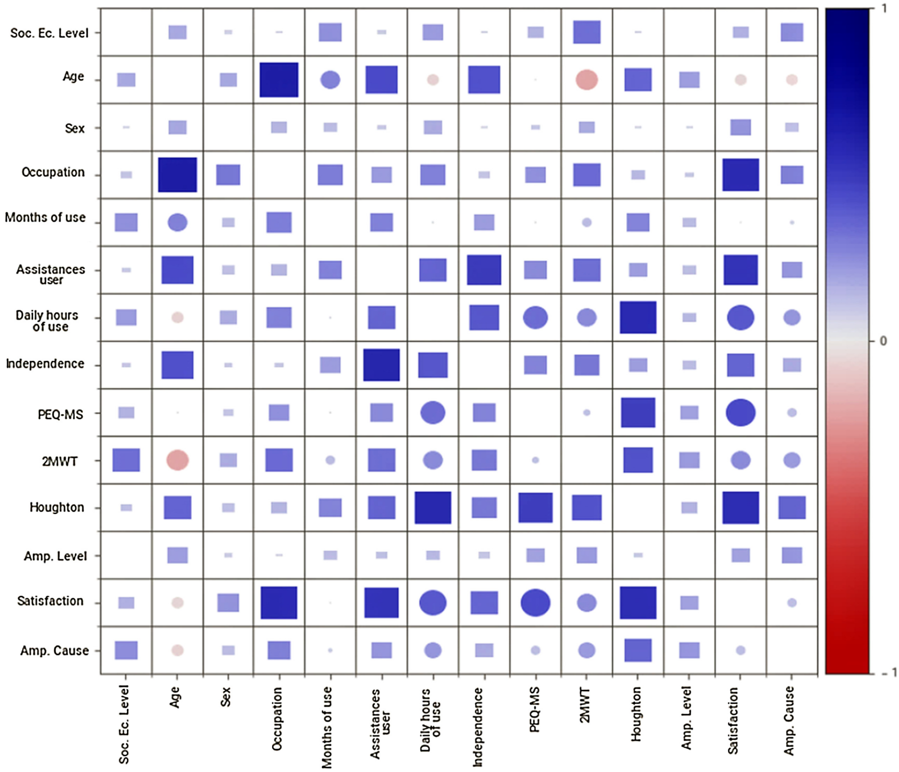

Fig. 1. Features associations using uncertainty coefficient and correlation ratio for categorical associations and Spearman correlation for numerical associations.

Associations between individual questions of the PEQ-MS were evaluated with the *"Self-perception of the ability to walk with the prosthesis"* and the results evidence strong correlations with specific aspects of the ambulation like: "Ability to walk in closed spaces using the prosthesis" (r = 0.70) and moderate correlations with the *"Ability to use a chair with a low seat"* (r = 0.48*), "Ability to go downstairs using the prosthesis"* (r = 0.47) and the *"Ability to walk on sidewalks and streets using the prosthesis"* (r = 0.45). The ANOVA test with the categorical variable of satisfaction as factor showed significance (p-value < = 0.05) for most of the questions on the PEQ-MS test, excluding the PEQJ and PEQH questions. Otherwise, the 2MWT and Houghton tests were not significant. Table 2 introduces the description of the individual questions of the PEQ-MS between the p-value resulting of the analysis of variance.

Table 2. Results of the analysis of variance using the satisfaction level as factor.

Feature	p-value	Description
PEQ-A	0.002	Self-perception of the ability to walk with the prosthesis
PEQ-B	0.019	Ability to walk in closed spaces using the prosthesis
PEQ-C	0.001	Ability to climb stairs using the prosthesis
PEQ-D	0.001	Ability to go downstairs using the prosthesis
PEQ-E	0.006	Ability to climb steep terrain using the prosthesis
PEQ-F	0.019	Ability to go down steep terrain using the prosthesis
PEQ-G	0.035	Ability to walk on sidewalks and streets using the device
PEQ-H	0.072	Ability to walk on wet surfaces using the prosthesis
PEQ-I	0.003	Ability to enter and exit a vehicle
PEQ-J	0.171	Ability to use a chair with an elevated seat
PEQ-K	0.004	Ability to use a chair with a low seat
PEQ-L	0.054	Ability to use a toilet
PEQ-M	0.006	Ability to take a bath
Occupation	0.077	Unemployed, employed, retired, house manager
Houghton	0.153	Houghton test score
2MWT	0.448	Measure in meters of the two-minute walk time test
Age	0.053	Age of the user
D. use time	0.051	Average hours of use per day

Moreover, MANCOVA analysis were useful to determine statistically significant differences between the adjusted means of the functional assessment test and scales, considering the error introduced by the covariates [31]. These results found a significant influence through the Wilks' Lambda test (p-value = 0.001).

Dimensional reduction analysis allowed the representation of the variation presented in the functional assessment tests and health-related behaviors, in a small number of factors. This step also remove multicollinearity in the features, making easier the modeling stage, and reducing the training time [32]. The 95% of the explained variance was reached using 17 factors from the original 30 features used in the analysis (PEQ-MS was used as the total average of the 13 questions). On the other hand, the 13 individual questions of PEQ-MS can be represented with 9 components maintaining the 95% of explained variance of data. Nonlinear transformations were applied using Uniform Manifold Approximation and Projection (UMAP), an algorithm for dimension reduction based on manifold learning and topological data analysis [23]. This method improved the compaction and separability of the different groups of patients as is evidenced in the Fig. 2, where the visualization of the UMAP embedding approach is compared with the projection using PCA.

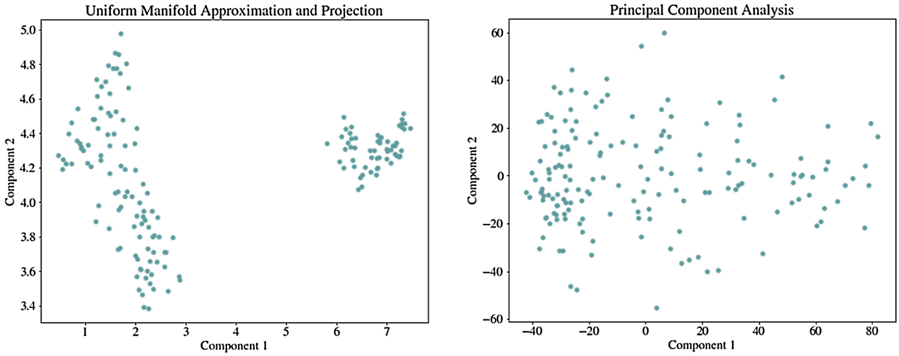

Fig. 2. Projection of the features in a 2D feature space with the UMAP embedding components (left) and the PCA components with the highest explained variance (right).

Agglomerative Hierarchical Clustering, K-means and Fuzzy C-means algorithms were trained with Python, using Scikit-learn, Scikit-fuzz and Pycaret packages. With the purpose of identify sets of clusters that are compact, with a small variance between members of the cluster, and well separated, we use 4 different index and coefficients for internal clusters validation. Silhouette coefficient and Davies Bouldin index suggest a conformation of 2 clusters in the dataset, and the Calinski-Harabasz and Dunn index suggest the conformation of 3 clusters. The results of these internal validation metrics used in 3 different clustering algorithms (Agglomerative Hierarchical Clustering (AC), K-means (KM) and Fuzzy C-means (FCM)) have been registered in the Table 3, where highlighted cells indicate the blue selection indicates the best metric found for the n-cluster, according to the criterion used by this. Only for Davies Bouldin index, a lower value will mean that the clusters are better.

Table 3. Internal validation results of Agglomerative Hierarchical Clustering (AC), K-means (KM) and Fuzzy C-means (FCM), with the Silhouette coefficient and Calinski Harabasz, Davies Bouldin and Dunn index.

Metric	Silhouette			Calinski-Harabasz			Davies Bouldin			Dunn		
Method	AC	KM	FCM	AC	KM	FCM	AC	KM	FCM	AC	KM	FCM
2 clust.	0.75	0.75	0.75	865.3	865.3	865.3	0.34	0.34	0.34	0.09	0.09	0.09
3 clust.	0.66	0.66	0.66	1246.7	1253.7	1253	0.46	0.46	0.46	0.13	0.13	0.13
4 clust.	0.61	0.62	0.62	1157.6	1176.8	1176	0.58	0.56	0.57	0.06	0.09	0.09
5 clust.	0.49	0.49	0.49	1202.7	1233.4	1229.4	0.71	0.71	0.73	0.11	0.11	0.11
6 clust.	0.42	0.50	0.51	1121.4	1169.2	846.5	0.86	0.73	0.88	0.08	0.07	0.04
7 clust.	0.42	0.43	0.38	1139.6	1168.2	1065.2	0.87	0.81	0.95	0.12	0.10	0.09
8 clust.	0.39	0.40	0.35	1115.7	1164.5	956.8	0.88	0.92	1.02	0.12	0.07	0.04

The visualization and interpretation stage were carried out using the mean decrease in impurity strategy with a random forest classifier fitted with the 3 groups generated by the Fuzzy C-means meth-od as labels. Satisfaction, PEQ-MS, Age, Daily use, and total adaptation time were the features with the highest relevance in the discrimination of the proposed groups, as the Fig. 3 shows. Fuzzy partition coefficient (FPC) was calculated and plotted in the figure to determine the quality of the cluster.

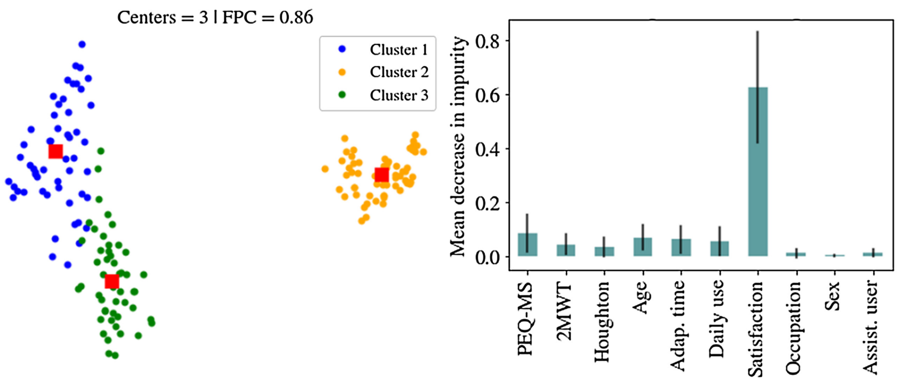

Fig. 3. Visualization and interpretation of the 3 clusters generated with the Fuzzy C-means algorithm in the UMAP projection, and the feature importance establish with the mean decrease impurity criterion.

4 Discussion

Considering the number of participants in similar works [2, 3] as well as relevance of the Mahavir Kmina corporation, which has over 3000 low-cost lower-limb prosthesis users in Latin America, we can assume that this study has an representative sample of the situation of amputees in Colombia [7].

Currently, there is no consensus of which variables and which reference value should be used to evaluate the adaptation and rehabilitation process in lower-limb prosthesis users. Moreover, there is a lack of methodologies that can propose these relevant factors and the reference values according to a specific study population. The associations found in the exploratory data analysis between the Houghton, PEQ-MS, and 2MWT tests are coherent with previous works; however, our results show that the strongest overall satisfaction was associated with the Houghton Scale as well as health-related behaviors such as the use of mobility assistances and the occupation of the user. Besides, the dimensional reduction analysis extended from the correlation analysis facilitates the understanding and modeling process of multidimensional problems such as the prosthesis adaptation process.

Consequently, the three clustering algorithms showed a similar trend in the internal evaluation metrics, and had three groups significantly separated from each other, as suggested by the Calinski-Harabasz and Dunn index value. Through the interpretation and visualization tools, these groups can be proposed as three states of adaptation in which a low-cost lower-limb prosthesis user can remain in their adaptation process with a high influence of the overall satisfaction. Although the determination of the appropriate number of groups for the problem is still on the frontier of knowledge, the possibility of identifying new states in patients through these methodologies allows for generating new tools in the monitoring and follow-up of people in the process of prosthetic rehabilitation. The modeling stage of this study is consistent with the Anderson behavioral model used in different health-rehabilitation processes, including prosthesis adaptation. This model can structure and organize healthcare management as a function of predisposing characteristics, establishing needs, and enabling resources. In this case, improving enabled resources and clinical strategies can impact the target population of these studies.

5 Conclusions

The user satisfaction is influenced by three general factors: Predisposing characteristics (e.g., the amputation level and the cause of amputation); Established need (e.g., age and occupation of the user); and enabling resources (e.g., the services offered to the users, like follow-up, maintenance, and training). To improve the results of the adaptation process, clinical policies should focus on user-centered fitting strategies supported by data science tools. Moreover, they should consider the development of monitoring systems that allow the use of prediction models, data collection for future researchers, and generation of automated reports and alerts in the evolution of prosthesis adaptation.

Although 199 participants are considered an acceptable representation considering the current state of the art, unsupervised learning algorithms can deliver more valuable results when trained with larger datasets. The development of software that integrates these components could provide remote diagnosis to patients in hard-to-reach areas and become a valuable source of data for specialists.

Acknowledgements. This work was supported by Medellín Higher Education Agency (Sapiencia) and the Ministry of Science, Technology, and Innovation of Colombia (Minciencias), and is part of a multi-disciplinary study called "Evaluation of a mixed strategy to improve the adherence of subjects amputated by Improvised Explosive Devices (IED) to the use of low-cost lower limb prostheses" with code 111580863475. We express our gratitude to the Mahavir Kmina Corporation and its entire work team, who were a great support throughout the research.

Conflict of Interest. The authors declare that there is no conflict of interest regarding the publication of this paper.

References

1. Arelekatti, V.N.M.: Detc2015–47385 Passive Prosthetic Knee for Users With Transfemoral Amputation, pp. 1–8 (2015)

2. Sansam, K., Neumann, V., O'Connor, R., Bhakta, B.: Predicting walking ability following lower limb amputation: a systematic review of the literature. J. Rehabil. Med. **41**(8), 593–603 (2009). https://doi.org/10.2340/16501977-0393
3. Baars, E.C., Schrier, E., Dijkstra, P.U., Geertzen, J.H.B.: Prosthesis satisfaction in lower limb amputees: A systematic review of associated factors and questionnaires. Medicine **97**(39), e12296 (2018). https://doi.org/10.1097/MD.0000000000012296
4. Zoellick, C.M.: Informe Mundial La Discapacidad Sobre R E S U M E N", B. R. Organ. Mund. la Salud, p. 27 (2011).
5. Biddiss, E.A., Chau, T.T.: Multivariate prediction of upper limb prosthesis acceptance or rejection. Disabil. Rehabil.: Assist. Technol. **3**(4), 181–192 (2008). https://doi.org/10.1080/17483100701869826
6. Salinas-Durán, F.A., et al.: "Guía de práctica clínica para el diagnóstico y tratamiento preoperatorio, intraoperatorio y posoperatorio de la persona amputada, la prescripción de la prótesis y la rehabilitación integral: Recomendaciones para el tratamiento de rehabilitación en adultos". Iatreia **29**(4), S-82–S-95 (2016)
7. Matamoros-Villegas, A., Plata-Contreras, J., Payares-Álvarez, K.: Correlation among tests and functional assessment scales in the follow-up of prosthetic adaptation in people with lower limb amputation. Rehabilitacion (2021)
8. Shamout, F., Zhu, T., Clifton, D.A.: Machine learning for clinical outcome prediction. IEEE Rev. Biomed. Eng. **14**, 116–126 (2021). https://doi.org/10.1109/RBME.2020.3007816
9. Chahar, R.: Computational decision support system in healthcare: a review and analysis. Int. J. Adv. Technol. Eng. Explor. **8**(75), 199–220 (2021). https://doi.org/10.19101/IJATEE.2020.762142
10. Ritter, M.A., Berend, M.E., Harty, L.D., Davis, K.E., Meding, J.B., Michael Keating, E.: Predicting range of motion after revision total knee arthroplasty. J. Arthro. **19**(3), 338–343 (2004). https://doi.org/10.1016/j.arth.2003.11.001
11. Syed Thouheed Ahmed, S., Thanuja, K., Guptha, N.S., Narasimha, S.: Telemedicine approach for remote patient monitoring system using smart phones with an economical hardware kit. In: 2016 International Conference on Computing Technologies and Intelligent Data Engineering. ICCTIDE 2016, pp. 1–4, 2016.
12. Zimina, E.Y., Novopashin, M.A., Shmid., A.V.:Application of medical data classification methods for a medical decision support system. CEUR Workshop Proceedinggs, vol. 2843 (2021)
13. Bose, E., Radhakrishnan, K.: Using unsupervised machine learning to identify subgroups among home health patients with heart failure using telehealth. CIN: Comput. Inf. Nurs. **36**(5), 242–248 (2018). https://doi.org/10.1097/CIN.0000000000000423
14. John, B., Wickramasinghe, N.: Clustering questions in healthcare social question answering based on design science theory. In: Wickramasinghe, N., Schaffer, J.L. (eds.) Theories to Inform Superior Health Informatics Research and Practice. HDIA, pp. 95–108. Springer, Cham (2018). https://doi.org/10.1007/978-3-319-72287-0_7
15. Nunes, M.A., Campos-Neto, I., Ferraz, L.C., Lima, C.A., Rocha, T.O., Rocha, T.F.: Adaptation to prostheses among patients with major lower-limb amputations and its association with sociodemographic and clinical data. Sao Paulo Med. J. **132**(2), 80–84 (2014). https://doi.org/10.1590/1516-3180.2014.1322572
16. Wong, C.K., Gibbs, W., Chen, E.S.: Use of the houghton scale to classify community and household walking ability in people with lower-limb amputation: criterion-related validity. Arch. Phys. Med. Rehabil. **97**(7), 1130–1136 (2016)
17. Houghton, A.D., Taylor, P.R., Thurlow, S., Rootes, E., McColl, I.: Success rates for rehabilitation of vascular amputees: implications for preoperative assessment and amputation level. British J. Surg. **79**(8), 753–755 (1992). https://doi.org/10.1002/bjs.1800790811

18. Sangve, S.S.B.S.M.: Clinical decision support system using SVM with the preservation of privacy. Int. J. Sci. Res **5**(7), 2122–2125 (2016)
19. Biddiss, E.A., Chau, T.T.: Multivariate prediction of upper limb prosthesis acceptance or rejection. Disabil. Rehabil.: Assist. Technol. **3**(4), 181–192 (2008). https://doi.org/10.1080/17483100701869826f
20. Kim, T.K.: Understanding one-way anova using conceptual figures. Korean J. Anesthesiol. **70**(1), 22–26 (2017)
21. Cardona, D., Uribe, J.: Identificación de las variables cinéticas, cinemáticas y funcionales en el proceso de adaptación protésica y la rehabilitación postprotésica en", Cent. Doc. Ing. (Bl. 20–146)Colección Tesis Electrónicas, Univ. Antioquia. (2018)
22. Malembaka, E.B., et al.: A new look at population health through the lenses of cognitive, functional and social disability clustering in eastern DR Congo: a community-based cross-sectional study. BMC Public Health **19**(1), 1–13 (2019)
23. McInnes, L., Healy, J., Melville, J.: UMAP: Uniform manifold approximation and projection for dimension reduction. arXiv (2018)
24. Indulska, M.: The curse of dimensionality in data quality. ACIS 2013 Proc. **2013**(December), 4–6 (2013)
25. Dilts, D., Khamalah, J., Plotkin, A.: Using cluster analysis for medical resource decision making. Med. Dec. Making **15**(4), 333–346 (1995). https://doi.org/10.1177/0272989X9501500404
26. Ogbuabor, G., Ugwoke, F.N.: Clustering algorithm for a healthcare dataset using silhouette score value. Int. J. Comput. Sci. Inf. Technol. **10**(2), 27–37 (2018). https://doi.org/10.5121/ijcsit.2018.10203
27. Graham, L.A., Fyfe, N.C.M.: Prosthetic rehabilitation of amputees aged over 90 is usually successful. Disabil. Rehabil. **24**(13), 700–701 (2002). https://doi.org/10.1080/09638280210142194
28. Windgassen, S., Moss-Morris, R., Goldsmith, K., Chalder, T.: The importance of cluster analysis for enhancing clinical practice: an example from irritable bowel syndrome. J. Mental Health **27**(2), 94–96 (2018). https://doi.org/10.1080/09638237.2018.1437615
29. Malli, S., Dr. Nagesh, H.R., Dr. Joshi, H.G.: A study on rural health care data sets using clustering algorithms. Int. J. Eng. Res. **3**(9), 546–548 (2014)
30. Granato, D., Santos, J.S., Escher, G.B., Ferreira, B.L., Maggio, R.M.: Use of principal component analysis (PCA) and hierarchical cluster analysis (HCA) for multivariate association between bioactive compounds and functional properties in foods: a critical perspective. Trends Food Sci. Technol. **72**, 83–90 (2018)
31. Siriwardena, G.J.A., Bertrand, P.V.: Factors influencing rehabilitation of arteriosclerotic lower limb amputees. J. Rehabil. Res. Dev. **28**(3), 35 (1991). https://doi.org/10.1682/JRRD.1991.07.0035
32. Islam, T., Rafa, S.R., Kibria, G.: Early Prediction of Heart Disease Using PCA and Hybrid Genetic Algorithm with k -Means, pp. 19–21 (2020)