Journal of
Modern Rehabilitation

October 2025, Volume 19, Number 4

Research Article

Logistic Regression Analysis of Functional Constipation Factors in the Elderly

Zeinab Shiravi[1*], Zahra Abdollahzade[2], Saeed Talebian[3], Fatemeh Alsadat Mirzadeh[4]

1. Department of Elderly Health, School of Rehabilitation, Tehran University of Medical Sciences, Tehran, Iran.
2. Department of Physiotherapy, Faculty of Paramedical and Rehabilitation Sciences, Mashhad University of Medical Sciences, Mashhad, Iran.
3. Department of Physiotherapy, School of Rehabilitation, Tehran University of Medical Sciences, Tehran, Iran.
4. Department of Geriatric Medicine, Ziaeian Hospital, Tehran University of Medical Sciences, Tehran, Iran.

Use your device to scan and read the article online

Citation Shiravi Z, Abdollahzade Z, Talebian S, Mirzadeh FA. Logistic Regression Analysis of Functional Constipation Factors in the Elderly. Journal of Modern Rehabilitation. 2025;19(4):415-426. http://dx.doi.org/10.18502/jmr.v19i4.19778

http://dx.doi.org/10.18502/jmr.v19i4.19778

Article info:
Received: 08 Mar 2025
Accepted: 03 Jun 2025
Available Online: 01 Oct 2025

Keywords:
Chronic constipation; Elderly; Supervised logistic machine learning

ABSTRACT

Introduction: Machine learning software programs are of great interest in medical sciences for their diagnostic and therapeutic applications. Elderly individuals can greatly benefit from these technologies due to their physical limitations. This study aimed to develop and evaluate a supervised machine learning model for predicting functional constipation (FC) in older adults.

Materials and Methods: Specific software was developed in Excel as a logistic regression supervised machine learning model (LR-SML 402). This software was developed based on a secondary analysis of existing data, including articles and doctoral dissertations on elderly individuals with FC who underwent colorectal evaluations using advanced laboratory equipment. The correlation between labeled data and colorectal parameter outputs was measured using 480 datasets from published sources and research laboratories. Strong correlations were observed between variables, such as age, body mass index, Wexner's questionnaire scores, and FC parameters.

Results: To validate the performance of LR-SML 402, the results were compared with those of a neural network model using SPSS software. The Excel-based model demonstrated strong performance in terms of sensitivity, specificity, and area under the curve.

Conclusion: The LR-SML 402 model shows that supervised machine learning using logistic regression may yield meaningful clinical predictions of FC indicators in the elderly. This approach can reduce diagnostic time and cost.

*** Corresponding Author:**
Zeinab Shiravi, Assistant Professor.
Address: *Department of Elderly Health, School of Rehabilitation, Tehran University of Medical Sciences, Tehran, Iran.*
Tel: *+98 (912) 3710970*
E-mail: *zshiravi@sina.tums.ac.ir*

Introduction

Artificial intelligence (AI) is one of the most important consequences of the development of extensive software technology in applied engineering. Today, this technique is ready to help patients, students, professors, and improve the treatment and health of society [1]. Recently, AI methods have been used to predict diseases and aging-related issues, enabling clinical professionals make decisions based on medical records. The improvement of AI, as one of the latest generations of modern technologies, has made rapid progress and plays a crucial role in predicting and classifying problems related to the elderly [2]. Many scientists in clinical and therapeutic research widely use this method for diagnosis, treatment, prediction, and improving healthcare effectiveness. Crucial and unprecedented developments are taking place in this field, especially in machine learning. Despite the development of machine learning, limitations still exist in the curricula of medical and rehabilitation schools worldwide, especially at the graduate level. Therefore, training and familiarization of academic staff, students, and other teaching staff with these technologies are essential [3].

However, with the increase in the number of older adults in society, which is expected to increase by 56% in the next 15 years (individuals over 60 years old) and the number of "elderly" (over 80 years old) will triple by 2050, sphincter-related problems will be one of the vital health issues and limitations facing the elderly.

Logistic regression supervised machine learning (LR-SML) offers good interpretability and low computational cost, making it a suitable classification algorithm for high-dimensional datasets. This is a statistical method that predicts the probability of an outcome based on one or more predictor variables [4].

Various machine learning approaches have been used to identify the onset of dementia and cognitive problems. Some studies have focused on the activities of daily living of the elderly to predict their cognitive level. Deep learning techniques can be used to detect anomalies in normal human behavior [5]. Ju utilized deep learning with brain network data and clinically relevant information (including the subject's age, sex, and *ApoE* gene) to construct a targeted auto-encoder network. This network successfully distinguished normal aging from mild cognitive impairment and early-stage Alzheimer's disease. The model presented by Ju is more stable and reliable than traditional methods and can help predict and prevent Alzheimer's disease in its early stages [6].

According to a review of studies conducted in this field, no research has focused on predicting the indicators of functional constipation (FC) in the elderly. This research aimed to design an LR-SML program based on AI to predict the rate and extent of FC in the elderly without performing time- consuming, and costly procedures.

Materials and Methods

Subjects

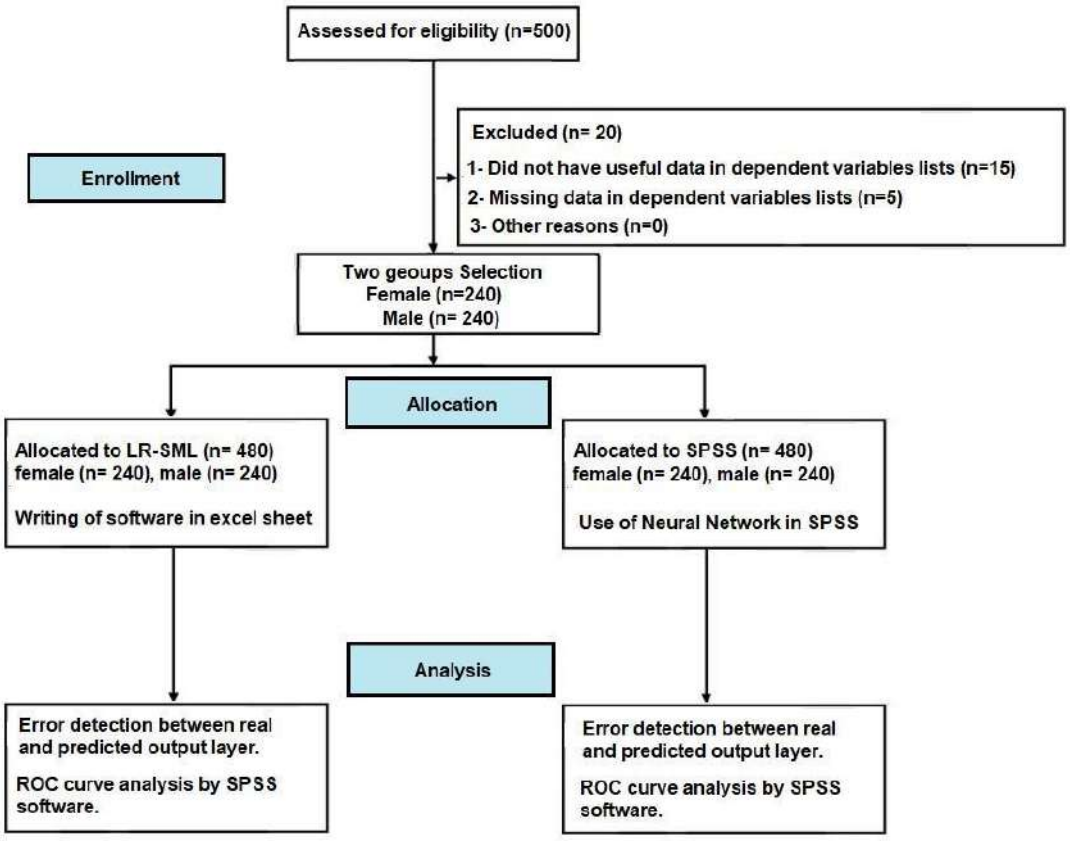
The prediction of FC values was performed in 480 cases, equally divided between male (n=240) and female (n=240) groups. Primary data (input layer) were collected from previous population studies (retrospective data from theses and clinics in Tehran). A consolidated standards of reporting trials (CONSORT) flow diagram illustrates the progression of participants from enrollment to analysis (Figure 1).

1. LR-SML

LR-SML was implemented to predict seven key anorectal physiological parameters: Resting anal pressure, rectal pressure during coughing, anal pressure during coughing, maximum pressure during squeezing, pressure during squeezing, threshold of anorectal inhibitory reflex, and defecation index. These output variables were analyzed about five input variables: Age, weight, height, body mass index (BMI), and Wexner questionnaire scores. The selection of these variables was based on established moderate to strong correlations documented in previous studies (Tables 1 and 2), which supported their predictive relationship with FC indicators in elderly populations. The LR-SML approach was chosen for its demonstrated effectiveness in handling such clinical prediction tasks while maintaining interpretability of results.

We used two independent software programs: 1) Logistic regression supervised machine learning in Excel, and 2) SPSS (neural network, multilayer perceptron). The sensitivity and specificity of the LR-SML software were evaluated and compared with the neural network

Multilayer perceptron in SPSS. Dedicated non-negative matrix factorization (NMF) equations were written in Excel to calculate multiple regression between input variables and predictions in the input layer. NMF is a technique in which a matrix V with nonnegative entries is factored into two matrices W and H (Figure 2).



JMR

Figure 1. CONSORT diagram comparing two machine learning software analyses

Matrix multiplication can be implemented by computing the column vectors of V as linear combinations of the column vectors in W using coefficients supplied by the columns of H. The relationship is expressed as: V=W·H

The equations for NMF are as follows:

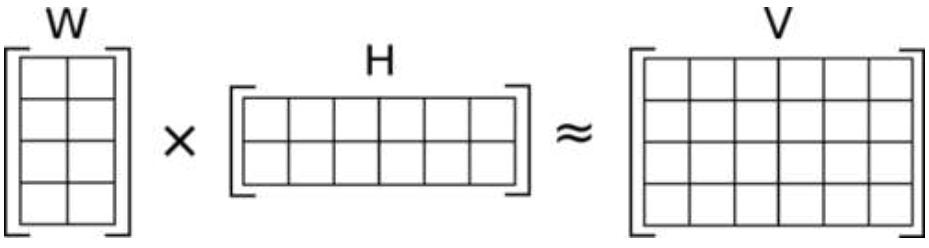
a. MINVERSE (MMULT (TRANSPOSE [A2:F241], A2:F241]),

where A2:F241 is the input layer (intercept "A," age "B," weight "C," height "D," BMI "E," and Wexner "F") for 240 cases in the Excel sheet (Figure 3).

b. MMULT (P2:U7, MMULT [TRANSPOSE [A2:F241], H2:N241]),

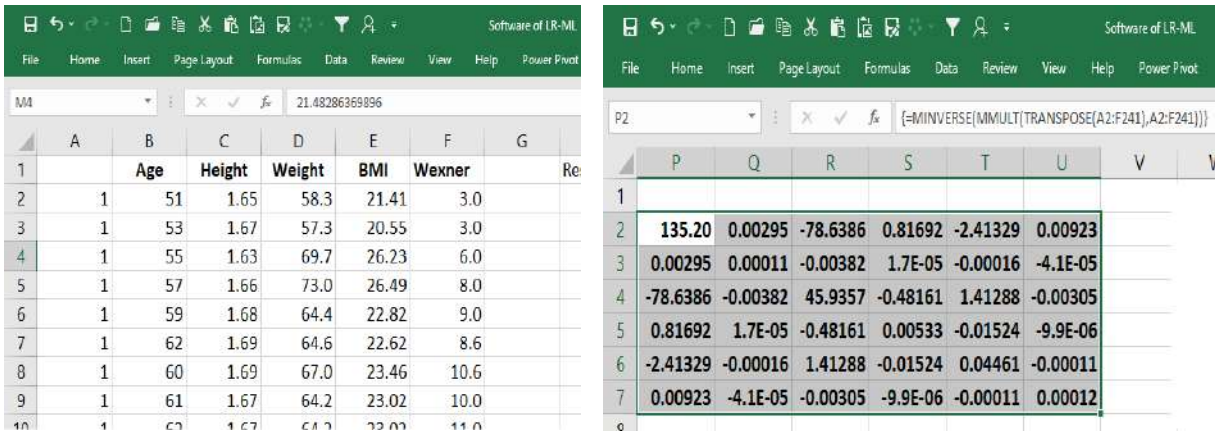
where P2:U7 contains the MINVERSE (MMULT [TRANSPOSE [intercept...]]) results and H2:N241 represents the primary output layer variables (resting anal pressure "H," rectal pressure during coughing "I," anal pressure during coughing "J," maximum pressure during squeezing "K," pressure during squeezing "L," threshold of anorectal inhibitory reflex "M," and defecation index "N") as implemented in the Excel spreadsheet (Figure 4).

c) MMULT ($A$2: $F$241, W2:W7 …AC7),



JMR

Figure 2. NMF diagram



JMR

Figure 3. Input layer variables (age, height, weight, BMI, Wexner) and NMF equation (MINVERSE [MMULT [TRANSPOSE [A2:F241], A2:F241]]) in Excel

where W2:W7 contains the multiple regression values (hidden layer) used to detect and predict the output layers corresponding to columns W, X, Y, Z, AA, AB, and AC. This equation predicts all output layer values at AE-AK columns in Excel sheet (Figure 5). Then, from this particular controlled matrix, in algorithm was used to determine the best estimate between inputs and outputs (Figure 6).

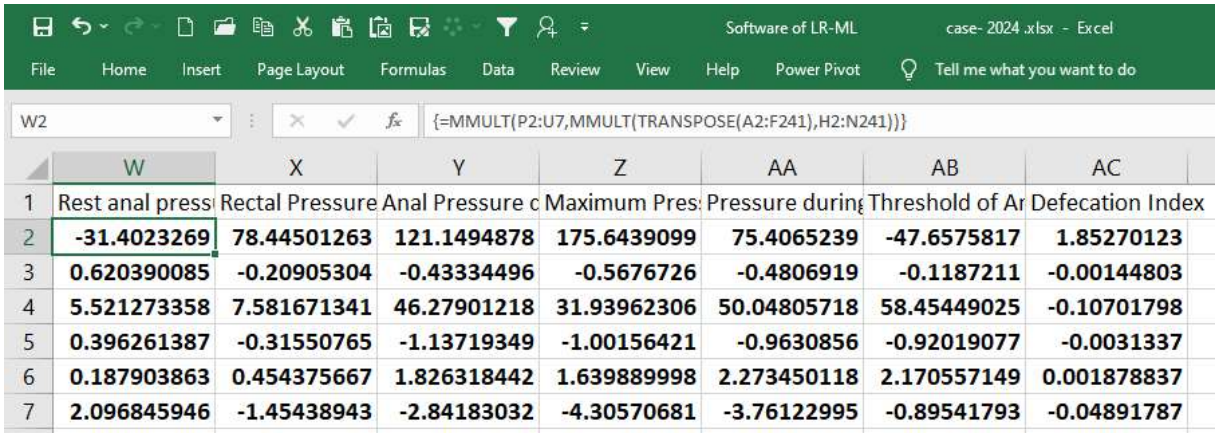
An optimization algorithm was then applied to the controlled matrix to determine the best estimate between the inputs and outputs.

2. Software under SPSS environment in the neural networks section

In SPSS software (all versions), the neural network module estimates output layers by evaluating the input layers and initial outputs. Using the multilayer perceptron option with a 70% training and 30% testing partition, we predicted output layers for FC. This allowed comparison with the LR-SML Excel implementation. The software generated: visualizations of hidden layers, calculation matrices with multiple regression, and predicted values stored in new columns (Figures 7 and 8).

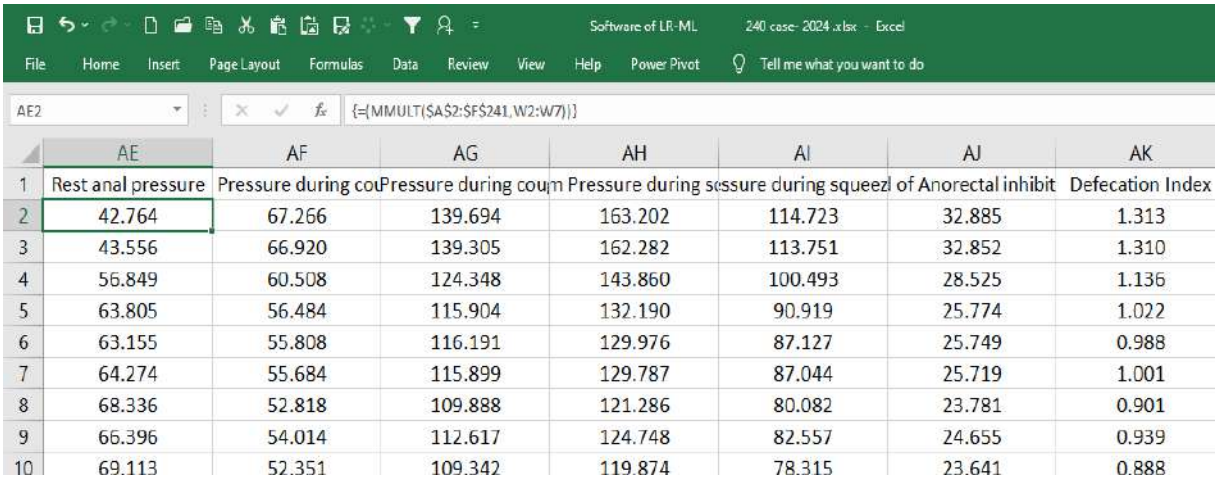
Results

As mentioned earlier, moderate to strong correlations were observed between the initial input and output. An appropriate algorithm was developed using supervised machine learning methods in the Excel environment. Various criteria were used to verify and validate the algorithm. A comparative analysis was performed between the radial basis function network machine learning algorithm and multiple logistic regression in SPSS, using labeled data for training and testing (Figure 9).



JMR

Figure 4. Hidden layer calculation using matrix multiplication: MMULT (P2:U7, MMULT [TRANSPOSE [A2:F241], H2:N241])



	AE	AF	AG	AH	AI	AJ	AK
1	Rest anal pressure	Pressure during cou	Pressure during coun	Pressure during s	essure during squeezi	of Anorectal inhibit	Defecation Index
2	42.764	67.266	139.694	163.202	114.723	32.885	1.313
3	43.556	66.920	139.305	162.282	113.751	32.852	1.310
4	56.849	60.508	124.348	143.860	100.493	28.525	1.136
5	63.805	56.484	115.904	132.190	90.919	25.774	1.022
6	63.155	55.808	116.191	129.976	87.127	25.749	0.988
7	64.274	55.684	115.899	129.787	87.044	25.719	1.001
8	68.336	52.818	109.888	121.286	80.082	23.781	0.901
9	66.396	54.014	112.617	124.748	82.557	24.655	0.939
10	69.113	52.351	109.342	119.874	78.315	23.641	0.888

JMR

Figure 5. Output layer calculation using matrix multiplication: MMULT ($A$2:$F$241, W2:W7) for columns AE to AK

The Mean±SD of the two groups (female and male) were measured (Tables 3 and 4).

The receiver operating characteristic (ROC) curve (SPSS software, version 22) was used to measure (TP, TN, FP, FN) and evaluate the specificity and sensitivity of the two algorithms with real variables from the primary data. The area under the curve was calculated using SPSS version 22 for both groups (Tables 5, 6 and 7; Figures 10 and 11).

Both analytical methods showed good predictive performance for most parameters. The accuracy percentage was calculated using Equation 1 [7]:

$$1.\text{Accuacy} = \frac{\text{Number of Correct Classifications}}{\text{Number of Total Classifications}} \times 100$$

The detection performance of the model was evaluated using two metrics: sensitivity and specificity, which reflect the model's ability to correctly identify positive and reject negative cases, respectively [8].

These metrics are defined as Equations 2 and 3:

$$2.\ \text{Sensitivity} = \frac{\text{TP}}{\text{TP TP} + \text{FN}}$$

$$3.\ \text{Specificity} = \frac{\text{TN}}{\text{TN TN} + \text{FP}}$$

Where TP (true positive) refers to the number of correctly predicted event cases, TN (true negative) to the number of correctly predicted non-event cases, FP (false positive) to the number of incorrectly predicted event cases, and FN (false negative) to the number of incorrectly predicted non-event cases.

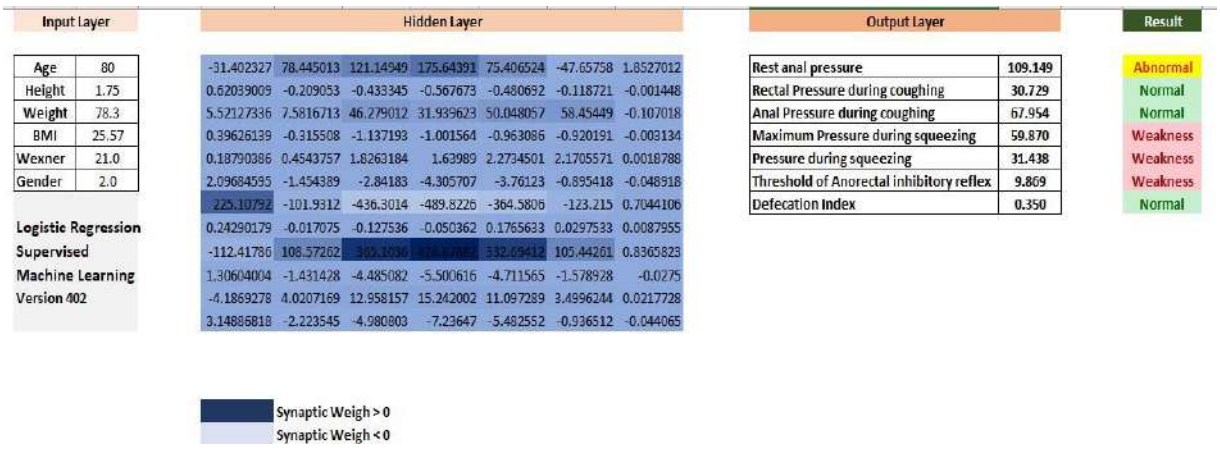


Figure 6. LR-SML software architecture

JMR

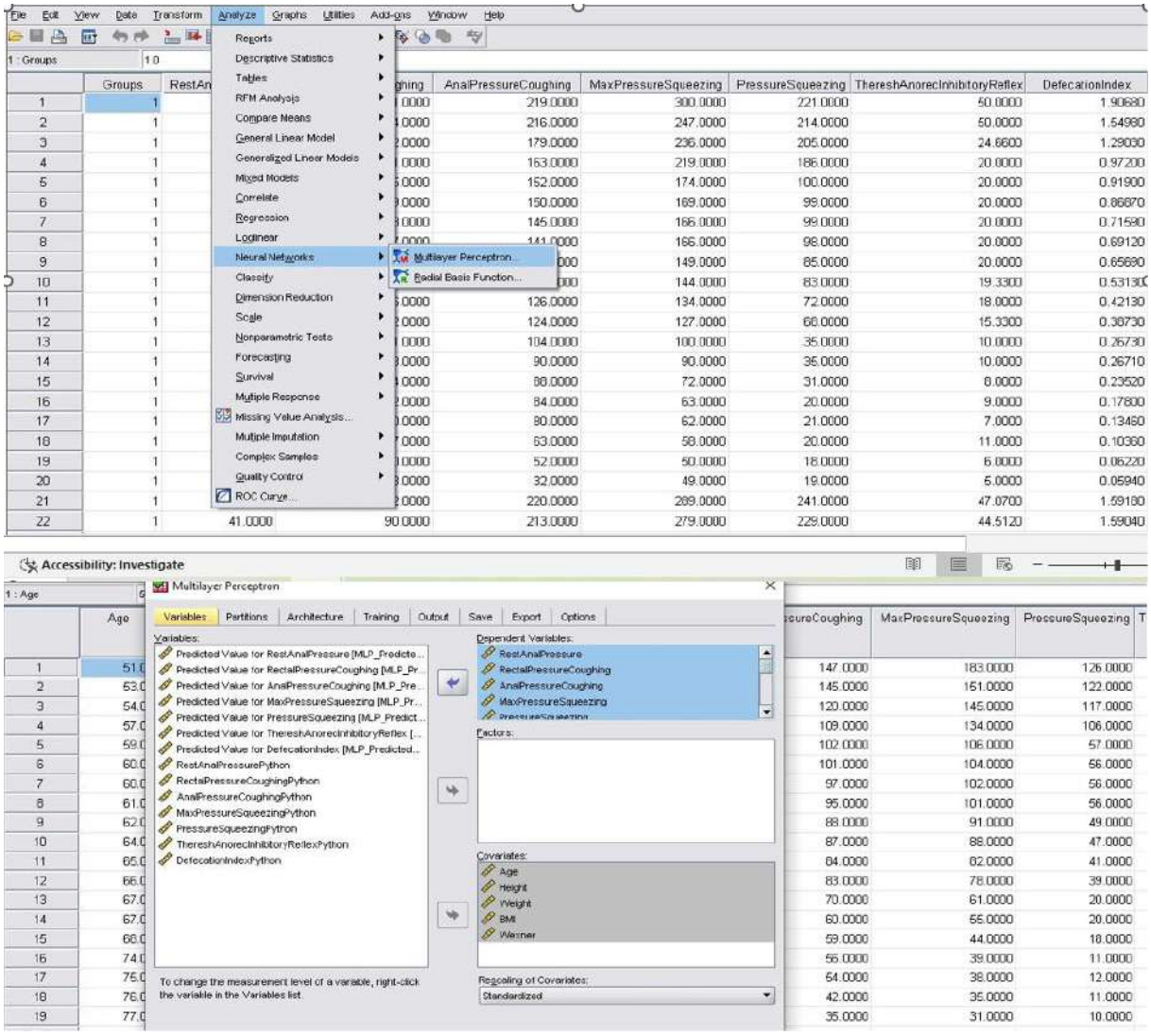


Figure 7. SPSS neural networks module: Configuration for output layer prediction from input and initial output layers

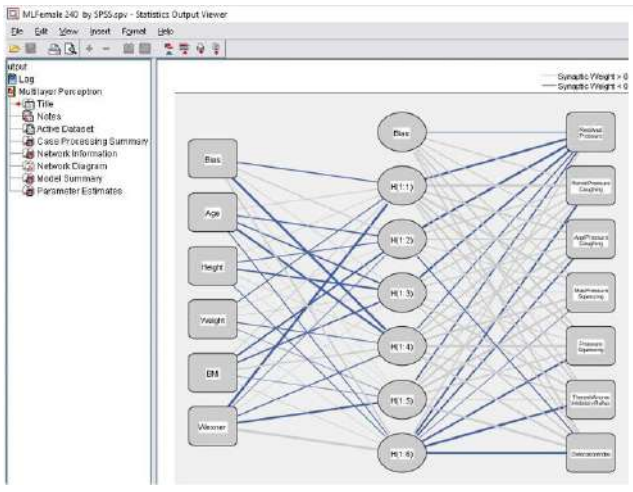


Figure 8. SPSS neural network output showing: (1) hidden layer neurons, (2) input layer interactions (excitatory/inhibitory), and (3) their combined effects on output layer predictions

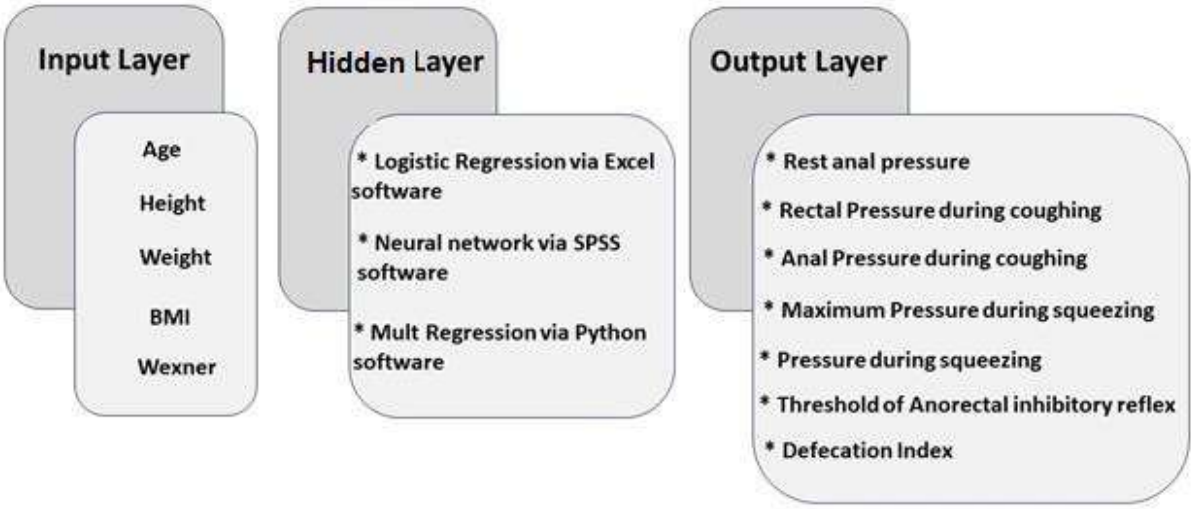


Figure 9. The three stages for the design of ML software

JMR

Table 1. Correlation between input layer variables and primary output layer variables in female participants (n=240)

Primary Output Layer	Input Layer				
	Age (y)	Height (m)	Weight (Kg)	BMI (Kg/m²)	Wexner
Rest anal pressure (mm Hg)	0.709	-0.222	0.691	0.631	0.916
Rectal pressure during coughing (mm Hg)	-0.653	0.182	-0.651	-0.579	-0.956
Anal pressure during coughing (mm Hg)	-0.664	0.192	-0.684	-0.607	-0.953
Maximum pressure during squeezing (mm Hg)	-0.648	0.181	-0.648	-0.576	-0.959
Pressure during squeezing (mm Hg)	-0.627	0.166	-0.61	-0.539	-0.945
Threshold of anorectal inhibitory reflex (cm³)	-0.629	0.14	-0.665	-0.560	-0.923
Defecation index (ratio)	-0.573	0.17	-0.594	-0.531	-0.944

JMR

Table 2. Correlations between input layer variables and output layer variables in male participants (n=240)

Primary Output Layer	Input Layer				
	Age (y)	Height (m)	Weight (Kg)	BMI (Kg/m²)	Wexner
Rest anal pressure (mm Hg)	0.661	-0.152	0.632	0.584	0.955
Rectal pressure during coughing (mm Hg)	-0.602	0.121	-0.637	-0.567	-0.914
Anal pressure during coughing (mm Hg)	-0.572	0.11	-0.597	-0.528	-0.858
Maximum pressure during squeezing (mm Hg)	-0.572	0.119	-0.611	-0.545	-0.865
Pressure during squeezing (mm Hg)	-0.553	0.137	-0.645	-0.582	-0.839
Threshold of anorectal inhibitory reflex (cm³)	-0.616	0.121	-0.792	-0.683	-0.893
Defecation index (ratio)	-0.539	0.165	-0.771	-0.699	-0.878

JMR

Table 3. Mean of 240 cases (female)

Parameter	Minimum	Maximum	Mean±SD
Age (y)	49	79.00	65.13±8.21
Height (m)	1.51	1.82	1.72±0.06
Weight (kg)	56.5	85.1	71.98±7.66
BMI (kg/m²)	17.95	36.18	24.38±3.61
Wexner questionnaire scale	2	30	17.24±7.79
Resting anal pressure (mm Hg)	43	176	87.82±23.36
Rectal pressure during coughing (mm Hg)	14	70	41.18±13.53
Anal pressure during coughing (mm Hg)	21	147	86.30±27.61
Maximum pressure during squeezing (mm Hg)	12	183	87.36±39.53
Pressure during squeezing (mm Hg)	1	137	51.55±33.96
Threshold of anorectal inhibitory reflex (cm³)	0	44	16.6±8.85
Defecation index (ratio)	0	1.7	0.55±0.42

JMR

As shown in Tables 6 and 7, both software systems demonstrated high true positive rates with good sensitivity across both groups, along with relatively good accuracy (Acc.). Sensitivity and specificity exhibit an inverse relationship; as sensitivity increases, specificity typically decreases, and vice versa. Highly sensitive tests tend to identify positive cases in patients with the condition, while highly specific tests effectively exclude the condition in unaffected patients. These findings suggest that both software tools can effectively predict FC difficulties in elderly populations.

Table 4. Mean of 240 cases (male)

Parameter	Minimum	Maximum	Mean±SD
Age (y)	49	79	65.13±8.2
Height (m)	1.51	1.87	1.74±0.07
Weight (kg)	54.5	90	74.93±8.36
BMI (kg/m²)	16.98	35.83	24.85±3.6
Wexner questionnaire scale	2	30	13.96±7.25
Resting anal pressure (mm Hg)	38	157	82.99±24.27
Rectal pressure during coughing (mm Hg)	7	92	47.61±17.84
Anal pressure during coughing (mm Hg)	18	220	107.43±42.58
Maximum pressure during squeezing (mm Hg)	0	300	115.3±62
Pressure during squeezing (mm Hg)	1	241	72.33±52.09
Threshold of anorectal inhibitory reflex (cm³)	0.07	50	17.88±9.73
Defecation index (ratio)	0.01	1.91	0.6±0.44

BMI: Body mass index.

JMR

Table 5. Area under the curve of two analytical methods across both groups

Parameter	Male Group		Female Group	
	SPSS	LR-SML	SPSS	LR-SML
Resting anal pressure (mm Hg)	0.69	0.799	0.5	0.491
Rectal pressure during coughing (mm Hg)	0.874	0.862	0.745	0.82
Anal pressure during coughing (mm Hg)	0.902	0.86	0.755	0.839
Maximum pressure during squeezing (mm Hg)	0.922	0.884	0.867	0.867
Pressure during squeezing (mm Hg)	0.861	0.865	0.556	0.882
Threshold of anorectal inhibitory reflex (cm^3)	0.858	0.88	0.859	0.867
Defecation index (ratio)	0.849	0.819	0.881	0.896

JMR

Table 6. Performance comparison between LR-SML and SPSS (neural network) models for female participants using ROC curve analysis

Parameter	LR-SML Model							SPSS Model						
	TP	TN	FP	FN	Sensitivity	Specificity	ACC	TP	TN	FP	FN	Sensitivity	Specificity	ACC
Resting anal pressure (mm Hg)	225	4	5	6	0.97	0.40	0.533	228	1	8	3	0.99	0.10	0.531
Rectal pressure during coughing (mm Hg)	216	7	8	9	0.96	0.44	0.581	226	2	8	4	0.98	0.18	0.573
Anal pressure during coughing (mm Hg)	215	8	7	10	0.96	0.50	0.575	223	5	5	7	0.97	0.45	0.567
Maximum pressure during squeezing (mm Hg)	212	5	16	7	0.97	0.23	0.678	215	7	9	9	0.96	0.41	0.670
Pressure during squeezing (mm Hg)	196	8	26	10	0.95	0.23	0.799	194	9	26	11	0.95	0.25	0.785
Threshold of anorectal inhibitory reflex (cm^3)	190	19	10	21	0.90	0.63	0.502	192	15	16	17	0.92	0.47	0.671
Defecation index (ratio)	178	16	28	18	0.91	0.36	0.822	187	16	19	18	0.91	0.44	0.824

Abbreviations: TP: True positive; TN: True negative; FP: False positive; FN: False negative; ACC: Accuracy.

JMR

Table 7. Performance comparison between LR-SML and SPSS (neural network) models for male participants using ROC curve analysis

Parameter	LR-SML Model							SPSS Model						
	TP	TN	FP	FN	Sensitivity	Specificity	ACC	TP	TN	FP	FN	Sensitivity	Specificity	ACC
Resting anal pressure (mm Hg)	225	4	5	6	0.97	0.4	0.515	233	0	5	2	0.99	0	0.519
Rectal pressure during coughing (mm Hg)	208	11	8	13	0.94	0.55	0.562	221	2	13	4	0.98	0.13	0.562
Anal pressure during coughing (mm Hg)	213	9	7	11	0.95	0.53	0.563	222	2	12	4	0.98	0.13	0.557
Maximum pressure during squeezing (mm Hg)	204	8	18	10	0.95	0.3	0.691	201	10	17	12	0.94	0.36	0.658
Pressure during squeezing (mm Hg)	199	10	19	12	0.94	0.33	0.790	202	11	14	13	0.94	0.42	0.792
Threshold of anorectal inhibitory reflex (cm^3)	189	18	13	20	0.9	0.56	0.500	182	22	12	24	0.88	0.63	0.683
Defecation index (ratio)	204	8	18	10	0.95	0.3	0.820	197	14	13	16	0.92	0.5	0.821

Abbreviations: TP: True positive; TN: True negative; FP: False positive; FN: False negative; ACC: Accuracy.

JMR

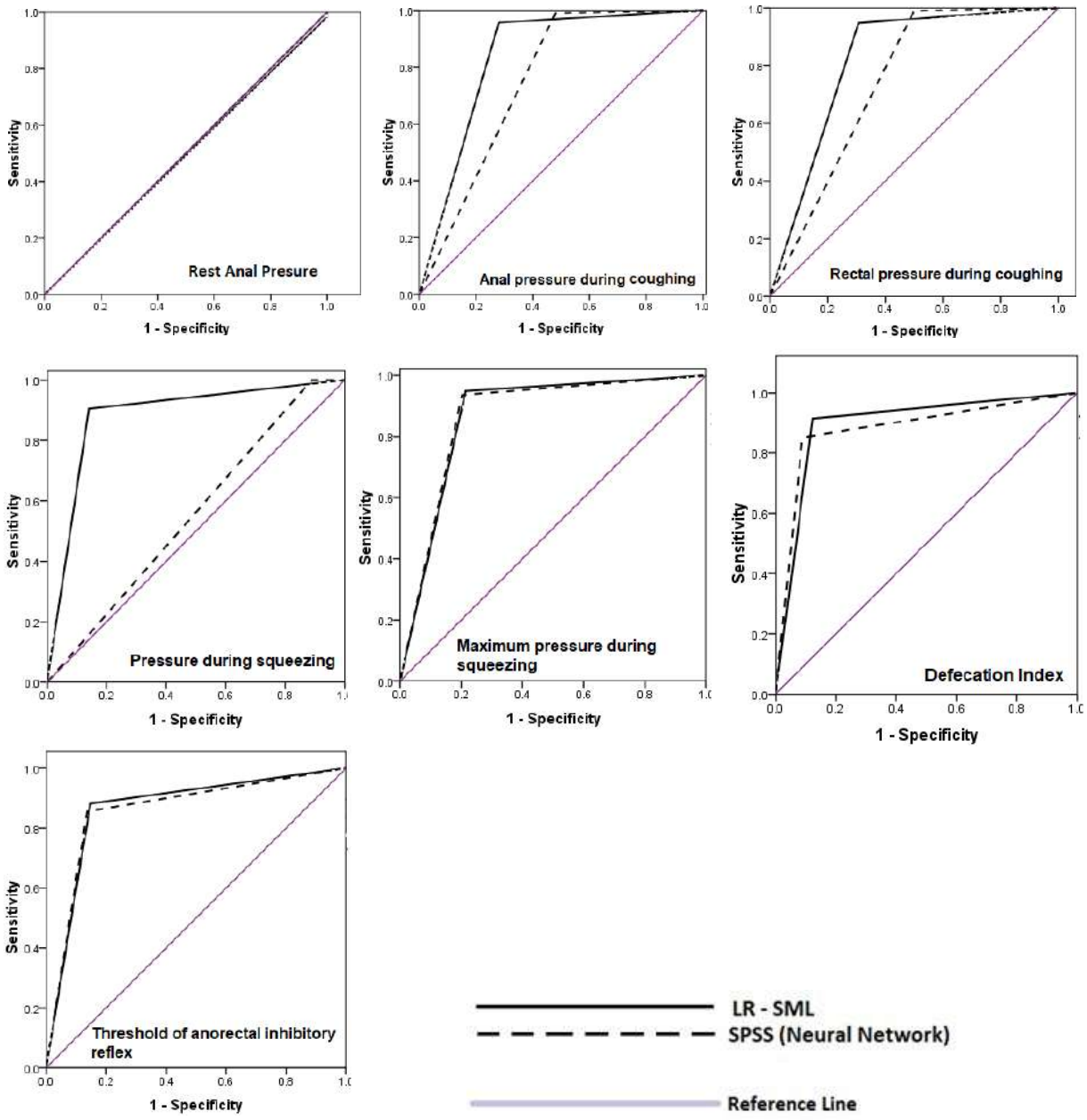


Figure 10. Sensitivity and 1-specificity comparison between two software programs for initial outcome prediction in female participants

Note: The LR-SML software demonstrated superior performance compared to the SPSS (neural network) software.

Discussion

This study employed a machine learning approach to develop a predictive model for estimating FC variables in elderly subjects using age, weight, height, BMI, and Wexner questionnaire scores, which demonstrated moderate to good correlations with outcome measures. Our findings indicate that the LR-SML software offers user-friendly operation and reliable predictive capability, consistent with previous applications in predicting brain wave activities in patients with tinnitus and estimating lumbar muscle activity in non-specific back pain [9, 10]. The model's sensitivity criteria effectively identified FC issues in the elderly population. While this study was limited to comparing LR-SML with SPSS (neural network), future research should include comparisons with other machine learning approaches, such as random forests. Designed specifically for clinicians and medical students, our model maintains simplicity as a key feature. Further external validation through multi-

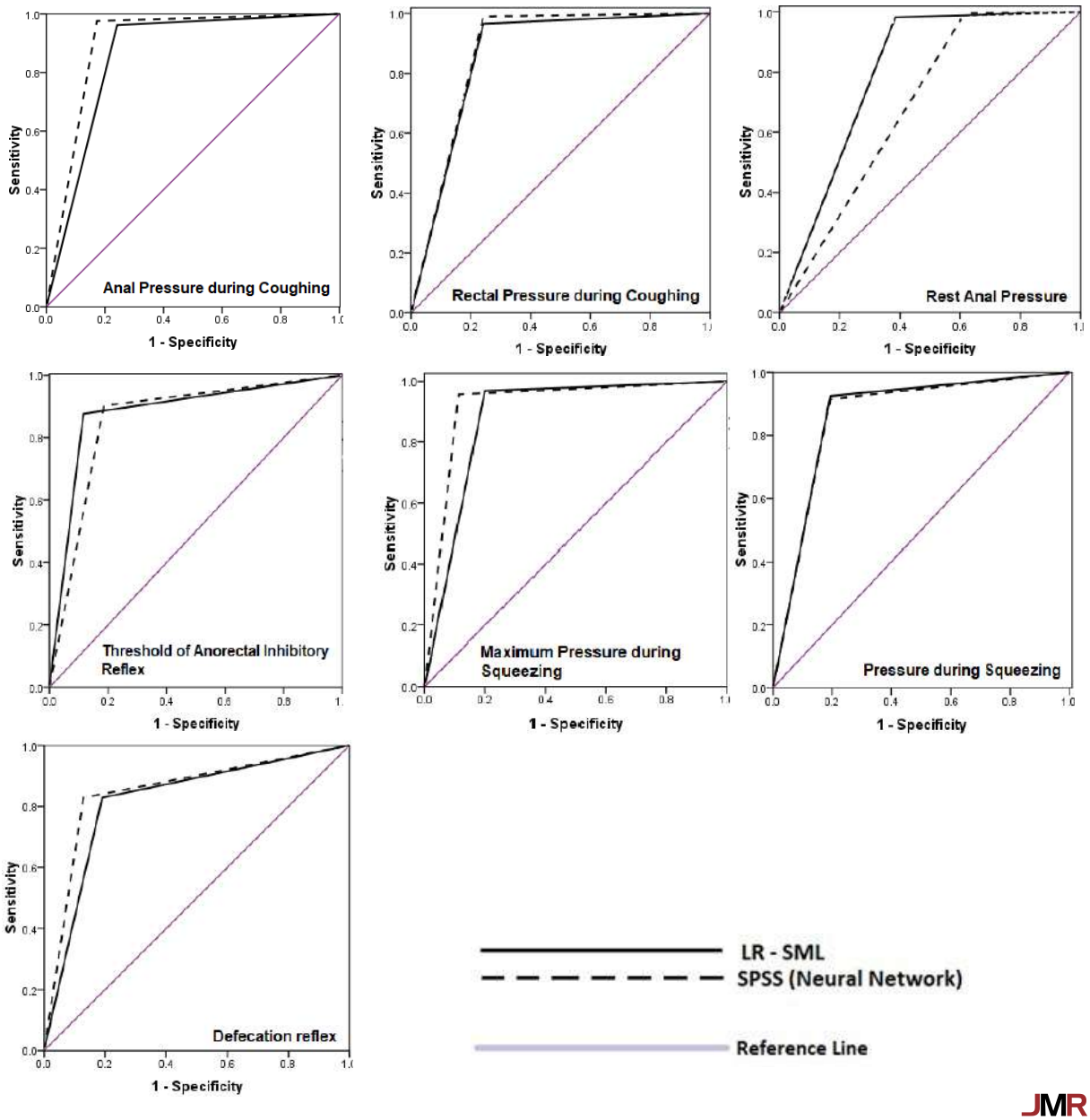


Figure 11. Comparison of sensitivity and 1-specificity changes between the two software programs for predicting initial outcomes in male participants

Note: The LR-SML software's performance was comparable to that of the SPSS (neural network) software.

step input data from diverse populations (both younger and older subjects) is needed to enhance and generalize this software.

Conclusion

The results demonstrate that LR-SML provides clinically relevant predictions of FC in elderly populations. The Excel-based implementation offers greater accessibility than other software options, and relative to SPSS (neural network), requires less prediction time without compromising outcome variables. Future studies should investigate the application of deep learning enhancements for other disabilities in this population. While this study focused exclusively on elderly participants with FC, we recommend additional investigations involving younger populations (both with and without FC) and elderly participants without constipation.

Ethical Considerations

Compliance with ethical guidelines

This study was approved by the Research Ethics Committees of Tehran University of Medical Sciences, Tehran, Iran (Code: IR.TUMS.MEDICINE.REC.1402.682).

Funding

This study was funded by Tehran University of Medical Sciences, Tehran, Iran.

Authors' contributions

Each author was involved in data analysis and contributed equally to the writing of the manuscript.

Conflict of interest

The authors declared no conflict of interest.

References

[1] Kolachalama VB, Garg PS. Machine learning and medical education. NPJ Digital Medicine. 2018; 1:54. [DOI:10.1038/s41746-018-0061-1] [PMID]

[2] Wassan JT, Zheng H, Wang H. Role of deep learning in predicting aging-related diseases: A scoping review. Cells. 2021; 10(11):2924. [DOI:10.3390/cells10112924] [PMID]

[3] Grunhut J, Wyatt AT, Marques O. Educating future physicians in artificial intelligence (AI): An integrative review and proposed changes. Journal of Medical Education and Curricular Development. 2021; 8:23821205211036836. [DOI:10.1177/23821205211036836] [PMID]

[4] Chen B. From logistic regression to deep learning: machine learning advances for credit card churn. International Journal of Enhanced Research in Management & Computer Applications. 2024; 13(2):31-5. [Link]

[5] Ishii H, Kimino K, Aljehani M, Ohe N, Inoue M. An early detection system for dementia using the M2 M/IoT platform. Procedia Computer Science. 2016; 96:1332-40. [DOI:10.1016/j.procs.2016.08.178]

[6] Ju R, Hu C, Li Q. Early diagnosis of Alzheimer's disease based on resting-state brain networks and deep learning. IEEE/ACM Transactions on Computational Biology and Bioinformatics. 2019; 16(1):244-57. [DOI:10.1109/TCBB.2017.2776910] [PMID]

[7] Rescio G, Leone A, Siciliano P. Supervised machine learning scheme for electromyography- Based pre-fall detection system. Expert Systems with Applications. 2018; 100:95-105. [DOI:10.1016/j.eswa.2018.01.047]

[8] Kutrani H, Eltalhi S, Ashleik N, editors. Predicting factors influencing survival of breast cancer patients using logistic regression of machine learning. The 7th International Conference on Engineering & MIS. 2021; 19:1-6. [DOI:10.1145/3492547.3492590]

[9] Mohsen S, Sadeghijam M, Talebian S, Pourbakht A. Use of some relevant parameters for primary prediction of brain activity in idiopathic tinnitus based on a machine learning application. Audiology & Neuro-otology. 2023; 28(6):446-57. [DOI:10.1159/000530811] [PMID]

[10] Salamat S, Montazeri V, Talebian S. Predicting trunk muscle activity in chronic low back pain: Development of a supervised machine learning model. Journal of Modern Rehabilitation. 2025; 19(2):148-63.[DOI:10.18502/jmr.v19i2.18348]