The myth of prostatic symptom scores: a look at the future

Wadie B. S., Badawi A.M.*, Ghoneim M. A.

Mansoura Urology & Nephrology Center, Mansoura University, Mansoura, Egypt

* Biomedical Engineering dept., Faculty of Engineering, Cairo University

Abstract:

INTRODUCTION: I-PSS is to be the prototype of symptom scores, which are exclusively used in the evaluation of porstatic patients and follow up of different treatment modalities. Many studies demonstrated poor or no correlation between BOO as diagnosed by pressure flow study and symptom severity as projected by I-PSS. This is the reason we tried the application of Artificial Neural Network (ANN) model in the evaluation patients with LUTS. PATIENTS AND METHODS: 460 patients were prospectively enrolled in this study, all of them had I-PSS, free flow PSA, TRUS and pressure flow plotting. ANN model (Prostatic Obstruction Predictor; POP) was designed. RESULTS: In the training set (305 patients), the model can predict obstruction in 94% (sensitivity of 94 % and specificity of 68 %). While in the test set (155), the model could predict obstruction in 87% of cases (sensitivity of 87% and specificity of 60%)

DISCUSSION: ANN is a relatively new modality in urologic diagnosis. Similar models were used in differentiation between benign and malignant prostatic enlargement as well in the diagnosis of prostatic adenocarcinoma. The accuracy of the model in the diagnosis of obstruction based on I-PSS is accepted, considering that statistical models

failed to demonstrate more than poor correlation between symptoms and objective obstruction. CONCLUSION: POP is an ANN model that helps in solving the conflict between symptoms and BOO; as objectively diagnosed by urodynamics.

INTRODUCTION:

I-PSS is a symptom score based on the AUA symptom index [1], which is designed for evaluation of males with LUTS due to benign prostatic hyperplasia (BPH).

Many studies relied on I-PSS as an objective tool used in the evaluation of patients with BPH and in the follow up after different treatment modalities [2, 3, 4]

Many studies failed to demonstrate a significant correlation between I-PSS and obstruction; as diagnosed by pressure flow studies [5, 6]

The use of Artificial Neural Network (ANN) was tried in the evaluation of patients with LUTS, aiming at optimizing the use of symptom scores.

PATIENTS AND METHODS:

Between 1997 and 1999, 520 patients were prospectively included in a protocol designed to evaluate the correlation of prostate symptom score with different objective parameters used in the evaluation of prostatic patients

Inclusion criteria are : males aged 45 years or more, having a total I-PSS score of 7 or more, and not being treated for BPH before.

Men with indwelling urethral catheter, associated urerthral stricture or bladder pathology i.e. malignancy or stone, as well as associated neuropathic bladder or major neurologic disease e.g. stroke, were excluded.

All patients had complete clinical examination, including DRE, and a neuro-urologic examination, serum total PSA, free flow rate, TRUS; sextant biopsy was done only if total PSA is over 4 ng/ml, the DRE is suspicious or the TRUS reveals suspicious nodules.

Out-patient cystoscopy, invasive urodynamic testing (in the form of filling and voiding water cystometry) are routinely carried out.

All patients answer I-PSS

The clinical examination included the DRE and the sensations in the saddle area, the anal tone and the bulbocavernosus reflex.

Serum PSA is requested. The test was done using the IMx technique (Abbott Lab., USA)

Free flow was carried out using a rotating drum flowmeter (Urodyn 1000, Dantec, Denmark). Only averaged value of three separate readings was considered.

Outpatient cystoscopy was carried out under local anaethesia (20 ml of 2 % Lidocaine gel instillation), using 17 F. rigid cystoscope.

The urethra was examined for strictures, the prostate for the type and degree of enlargement and the bladder for associated pathology.

Urodynamic testing was carried out using a multichannel computerized system (Duet, Dantec, Denmark).

Examination data are stored as database files, with a modification of the built-in database of the system.

Filing water cystometry was done using 8 F. dual urodynamic catheter with terminal hole and one side hole (Porges, France).

The initial part of the test was measurement of the post –voiding residua urine (PVR); measured within 5 minutes of the patient's last free void; using catheter.

Rate of filling is 50 ml/min, sometimes increased to 100 ml/min, if there is an evidence of low amplitude uninhibited contractions.

This is followed by voiding cystometry in erect position, with the transurethral catheter in place.

Technique and specifications of the urodynamic testing was conforming to the ICS standardization.[7]

Results are filtered and saved with interpretation of the pressure flow pattern with the use of the linear Passive Urethral Resistance Relation (Lin PURR) as described by Schafer [8]

Grades 0 and 1 are considered non- obstructed, grade 2 is equivocal and grades 3,4,5and 6 are considered obstructed in ascending manner[9]

On the same day of the urodynamic testing, the patient was asked to answer a standardized, validated Arabic form of the I-PSS questionnaire.

Four hundred and sixty patients were evaluable.

Data of the patients were fed to an ANN specially designed to accomplish this task, using MATLAB software (Mathworks, USA).

ANN is a complex computational system capable of undertaking a large number of complex mental tasks[10].

In doing so, it mimics the functions of the human brain, hence the name " artificial intelligence"

A Multilayer Perceptron (MLP), in its simplest description, is a network consisting of series of processing elements (neurons) arranged in layers. Each of these neurons is capable of simple computational processes, data are being presented in the back –propagation model to different neurons for a large number of times (epochs or iterations)[10]

In the POP model, a back propagation MLP is established

Unsupervised learning, using the K- means and fuzzy logic principles was first done.

This is followed by back propagation supervised learning.

662,000 iterations were needed to develop a model that has the lowest obtainable mean square error

The network consists of 8 neurons in the input layer, 25 neurons in the hidden layer and 3 neurons in the output layer.

Figure 1 shows the arbitrary description of a back propagation MLP

Training set consists of the records of 305 patients and the testing set consists of the records of 155 patients.

RESULTS:

According to the symptom severity, 38 patients (8.3%) were found to have mild symptoms, 241 (52.4%) moderate symptoms and 180 (39.1%) severe symptoms

The mean values for the patients' answers of the individual questions are demonstrated in table 1

The mean value of the individual question scores range from 2.03 to 3.03 and the mean value of the quality of life index is 4.4 "Unhappy"

The mean value of the total score as well as the voiding (obstructive) score (the sum of questions 1,3,5,and 6) and the storage (irritative) score (the sum of questions 2, 4 and 7) are demonstrated in table 1.

The output of the model is classified into obstructed, equivocal and non-Obstructed.

Although the original pressure flow analysis was interpreted in terms of LinPURR, which is a seven-band nomogram, yet, for simplicity, a categorical classification of the nomogram was used (grades 0 and 1 are non- obstructed, grade 2 is equivocal and grades 3, 4, 5, and 6 are considered obstructed) [8, 9]

The overall number of patients having Schafer grade of 3 or more is 285 (60%), grade 2 (considered to be equivocal) is 88 (18%), and grade 0 and

1 are 85 (22%). Table 2 demonstrates the distribution of different grades of obstruction among the training and testing groups.

In the training set, the sensitivity and specificity of the model to predict obstruction are 94% and 68%.

Sensitivity and specificity to predict non-obstruction are 68% and 85%; sensitivity and specificity to predict equivocal cases are 56% and 86% respectively.

Mean sensitivity of the model in training set is 72.7% and specificity is 79%

Table 3 shows the confusion matrix of the training set

While in the test set, the sensitivity and specificity of the model to predict obstruction are 87% and 40%, sensitivity and specificity to predict non-obstruction are 60% and 82%, sensitivity and specificity to predict equivocal cases are 49% and 83% respectively.

Mean sensitivity of the model in testing set is 65.3% and specificity is 68.3%.

Confusion matrix of this group is demonstrated in table 4

DISCUSSION:

Although it is used by many urologists to evaluate treatment outcome of BPH, I-PSS did not correlate to objective parameters customarily used in the evaluation of BPH [11] [12]

To date, the gold standard of the diagnosis of BOO is pressure flow studies [13]

Based on conventional statistics, it was demonstrated that correlation between I-PSS and BOO as diagnosed by pressure flow studies, is approaching zero.[14]

The method of LinPURR was considered the standard for analysis of the pressure flow plots, as it is considered more sensitive than other available methods in the quantification of obstruction.[15]

Besides, the size of the equivocal zone (grade II) is comparable to the ICS method, which is smaller than other comparable nomograms [9]

In an earlier report, statistical analysis revealed no correlation between the severity of symptoms and BOO, as rated by pressure flow study.

This is the rational behind using ANN to verify if there could be an improvement in diagnostic yield of the I-PSS.

In the POP model, a back propagation MLP is established

In medicine, ANN was exclusively used in decision-making and classification systems in different fields.

Fuzzy logic was used to differentiate between cirrhotic liver and normal or fatty. An eight dimensional vector was fed to this model and the output was one of the three liver patterns [16]

In urology, the diagnosis and prognostication of prostatic adenocarcinoma had a considerable share of ANN applications[17] [18]

One of the well-known applications of ANN, the ProstAsure Index (PI), was well studied in many publications.

In the original paper, Stamey et al described a MLP model capable of differentiating benign from malignant prostate with a sensitivity of 815 and a specificity of 92%

The index depends on 4 input parameters (patient's age, serum PSA, prostatic acid phosphatase, and total creatine kinase)

The output consists of a mathematical scale; 0 or less, 0.1-0.5, 0.5-1 and 1 or more which correspond to normal, BPH, suspicious of malignancy and mostly malignant respectively. [19]

Our results demonstrate that, using ANN, I-PSS could be used in the diagnosis of BOO with an overall sensitivity of 65% and a specificity of 68%.

The lowest sensitivity is encountered with equivocal cases.

Among our patients, 19% are classified by the LinPURR to be "equivocal"

Since all pressure flow nomogram have an equivocal zone [9, 12], this may explain the relatively low overall sensitivity of the model, as the sensitivity in the equivocal zone is only 49%.

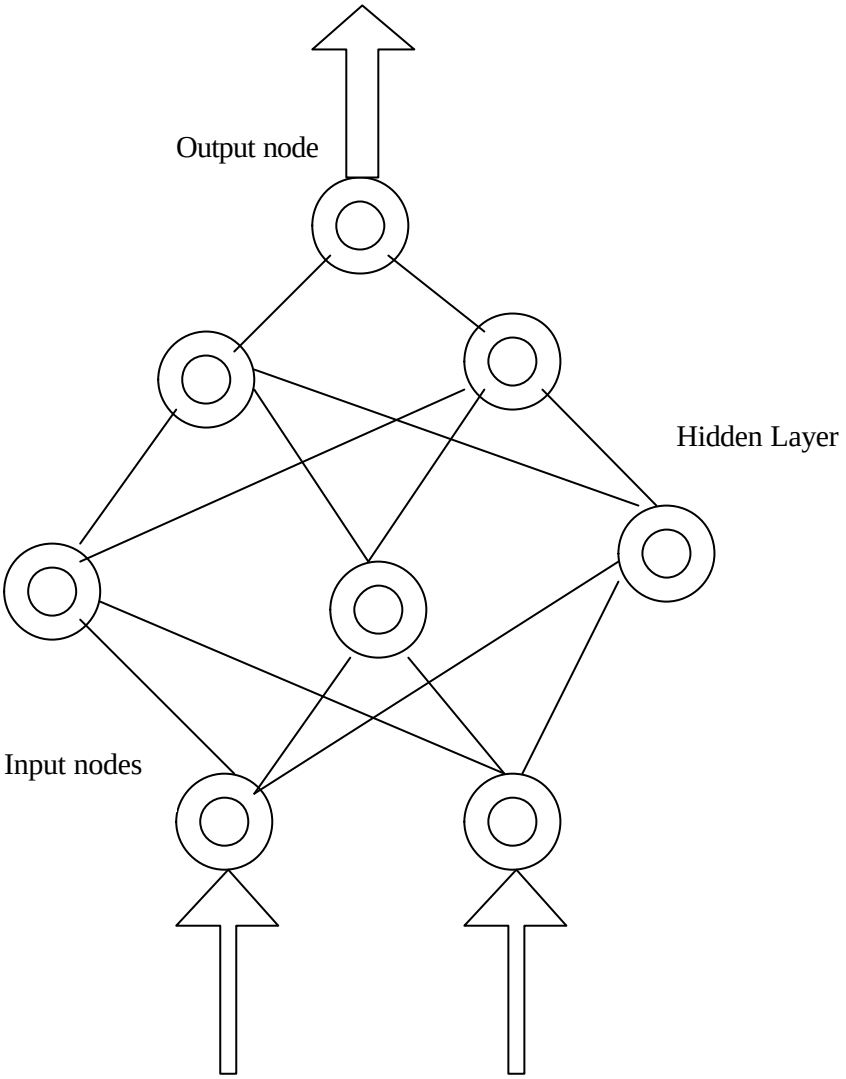
Compared to ordinary statistical regression models, the POP model is a step forward in the way to diagnose BOO based on symptom scores.

CONCLUSION:

POP model is a helpful tool in objectifying non-nominal symptom score. The utilization of the ANN's ability of making decision boundaries from nonlinear data has made possible the evolution of a fairly reliable diagnostic tool based on symptoms.

Figure 1:

The architecture of a back-propagation MLP model



A simple back-propagation neural network

Table 1:

Mean value of the score of individual questions

Question	**Mean score +/- Standard deviation**
1-Incomplete emptying	2.42 +/- 1.95
2-Frequency	2.29 +/- 2.03
3- Intermittency	2.68 +/- 1.81
4- Urge	2.09 +/- 1.94
5- Weak stream	3.03 +/- 1.63
6- Hesitancy	2.03 +/- 1.84
7- Nocturia	3.00 +/- 1.42
8- quality of life index	4.40 +/- 1.49
Total score (35)	17.53 +/- 6.65
*Score of obstructive symptoms (20)	10.15 +/-4.58
§ Score of irritative symptoms (15)	7.38 +/-3.98

**Sum of questions1, 3, 5&6, total is 20*
§ Sum of questions 2,4 &7, total is 15

Table 2:

Number of cases

	Schafer's grade	Training set	Testing set
Non-obstructed	0	10	6
	1	46	23
Equivocal	2	58	30
Obstructed	3	69	34
	4	67	34
	5	40	20
	6	15	6
Total (458):		305	153

Non-obstructed: **Equivocal:** **Obstructed:**

The distribution of different LinPURR grades among training and testing sets

Table 3:

Obstruction (projected by the POP)		Obstruction (determined by urodynamics)		
		Non-obstructed	Equivocal	Obstructed
	Non-obstructed	68	10	22
	Equivocal	13	56	31
	Obstructed	2	4	94

Training set, Confusion matrix %
(n= 305)

Table 4:

Obstruction (determined by urodynamics)

Obstruction (projected by POP)	Non-obstructed	Equivocal	Obstructed
Non-obstructed	60	8	32
Equivocal	14	49	37
Obstructed	4	9	87

Testing set, confusion matrix %
(n = 153)

REFERRENCES:

1 Barry, M. J., Fowler, F. J. Jr., O'Leary, M. P., Bruskewitz, R. C., Holtgrewe, H. L., Mebust, W. K., Cockett, A.T.: The American Urological Association symptom index for benign prostatic hyperplasia. J. Urol. 148: 1549, 1992.

2 Hakenberg, O. W., Pinnock, C. B., Marshall, V. R.: Does evaluation with the International Prostate symptom Score predict the outcome of transurethral resection of the prostate? J. Urol. 158:94, 1997.

3 Arai, Y., Okubo, K., Okada, T., Maekawa, S., Aoki, Y., Maeda, H.: Interstitial laser coagulation for management of benign prostatic hyperplasia: a Japanese experience. J. Urol. 159: 1961, 1998.

4 Stoner, E., and the finasteride study group: the clinical effects of a 5 alpha-reductase inhibitor, finastride, on benign prostatic hyperplasia. The finasteride study group. J. Urol. 147:1298, 1992.

5 Bosch, J. L. H. R., Hop, W. C. J., Kirkels, W. J., Schroder, F. H.: The International Prostate Symptom Score in a community- based sample of men between 55 and 74 years of age: Prevalence and correlation of

symptoms with age, prostate volume, flow rate, and residual urine volume. Br. J. Urol. 75: 622, 1995.

6 Ko D. S. C., Fenster H.N., Chambers K., Sullivan L., Jens M., Goldenberg S.L.: The correlation of multichannel urodynamic pressure-flow studies and the American Urological Assocation symptom ndex in the evaluation of benign prostaic hyperplasia J. Urol. 154: 396-398, 1995

7 Abrams P., Blaivas J. B., Stanton S. L., Andersen J. T.: The standardization of terminology of lower urinary tract function W. J.Urol.6: 233, 1989.

8 Schafer, W.: Analysis of bladder-outlet function with the linearized passive urethral resistance relation, linPURR, and a disease-specific approach for grading obstruction: from complex to simple. W. J. Urol. 13:47, 1995.

9 Grifiths D.J., Hofner K., van Mastrigt R., Rollema H.J., Spangberg A., Gleason D.: Standardization of Terminology of Lower Urinary Tract: pressure- flow studies of voiding, Urethral resistance and urethral Obstruction. Neurourol. Urodyn.16: 1, 1997.

10 Armoni A.: Use of Neural Networks in Medical Diagnosis. M.D. Computing. 15, No.2: 100, 1998.

11 Yalla, S. Y., Sullivan, M. P., Lecamwasam, H. S., Du Beau, C. E., Vickers, M. A., Cravalho, E. G.: Correlation of the American Urological Association symptom index with obstructive and nonobstructive prostatism. J.Urol. 153:674, 1995.

12 Witjes W.P., Aarnick R. G., Ezz el Din K., Wijkstra H., Debruyne F. M., de La Rosette J.J.: The correlation between prostate volume, transition zone volume, transition zone index and clinical and urodynamic investigations in patients with lower urinary tract symptoms. Br. J Urol. 80: 84, 1997.

13 Griffiths, D. J.: Pressure- flow studies of micturition. Urol. Clin. N. Am., 23: 279, 1996.

14 Madersbacher S., Pycha A., Klingler C. H., Schatzl G., Marbergher M.: The international prostate symptom score in both sexes: a urodynamic-based comparison. Neurourol. Urodyn.18: 173, 1999.

15 Khoury, J. M., Marson, L., Carson, C. C.: A comparative study of the Abrams-Griffiths nomogram and the linear passive urethral resistance relation to determine bladder outlet obstruction. J. Urol. 159: 758, 1998.

16 Badawi A. M., Derbala A. S., Youssef A.M.: Fuzzy logic algorithm for quantitative tissue characterization of diffuse liver diseases from ultrasound images. Int. J. Medical Informatics 55:135, 1999.

17 Snow P.B., Smith D.S., Catalona W.J.: Artificial neural networks in the diagnosis and prognosis of prostate cancer: a pilot study J.Urol. 152:1923, 1994

18 Douglas T., Connelly R., McLeod D. et al: Neural network analysis of pre-operative and post-operative variables to predict pathologic stage and recurrence. J.Urol. 155:487, 1996

19 Stamey T.A., Barnhill S. D., Zhang Z. et al: Effectiveness of ProstAsure TM in detecting prostate cancer (PCa) and benign prostatic hyperplasia (BPH) in men age 50 and older. J. Urol. 155:436, 1996