

Validation of artificial neural network models for predicting biochemical markers associated with male infertility

A. S. Vickram, A. Rao Kamini, Raja Das, M. Ramesh Pathy, R. Parameswari, K. Archana & T. B. Sridharan

To cite this article: A. S. Vickram, A. Rao Kamini, Raja Das, M. Ramesh Pathy, R. Parameswari, K. Archana & T. B. Sridharan (2016): Validation of artificial neural network models for predicting biochemical markers associated with male infertility, Systems Biology in Reproductive Medicine

To link to this article: <http://dx.doi.org/10.1080/19396368.2016.1185654>



Published online: 21 Jun 2016.



Submit your article to this journal [↗](#)



View related articles [↗](#)



View Crossmark data [↗](#)

RESEARCH ARTICLE

Validation of artificial neural network models for predicting biochemical markers associated with male infertility

A. S. Vickram^a, A. Rao Kamini^b, Raja Das^c, M. Ramesh Pathy^d, R. Parameswari^d, K. Archana^d, and T. B. Sridharan^d

^aIndustrial Biotechnology Division, School of Biosciences and Technology, VIT University, Tamil Nadu, India; ^bBangalore Assisted Conception Centre Healthcare Pvt. Ltd., Andrology, Bangalore, Karnataka, India; ^cSchool of Advanced Sciences, VIT University, Tamil Nadu, India; ^dSchool of Biosciences and Technology, VIT University, Tamil Nadu, India

ABSTRACT

Seminal fluid is the secretion from many glands comprised of several organic and inorganic compounds including free amino acids, proteins, fructose, glucosidase, zinc, and other scavenging elements like Mg^{2+} , Ca^{2+} , K^+ , and Na^+ . Therefore, in the view of development of novel approaches and proper diagnosis to male infertility, overall understanding of the biochemical and molecular composition and its role in regulation of sperm quality is highly desirable. Perhaps this can be achieved through artificial intelligence. This study was aimed to elucidate and predict various biochemical markers present in human seminal plasma with three different neural network models. A total of 177 semen samples were collected for this research (both fertile and infertile samples) and immediately processed to prepare a semen analysis report, based on the protocol of the World Health Organization (WHO [2010]). The semen samples were then categorized into oligoasthenospermia ($n=35$), asthenospermia ($n=35$), azoospermia ($n=22$), normospermia ($n=34$), oligospermia ($n=34$), and control ($n=17$). The major biochemical parameters like total protein content, fructose, glucosidase, and zinc content were elucidated by standard protocols. All the biochemical markers were predicted by using three different artificial neural network (ANN) models with semen parameters as inputs. Of the three models, the back propagation neural network model (BPNN) yielded the best results with mean absolute error 0.025, -0.080, 0.166, and -0.057 for protein, fructose, glucosidase, and zinc, respectively. This suggests that BPNN can be used to predict biochemical parameters for the proper diagnosis of male infertility in assisted reproductive technology (ART) centres.

Abbreviations: AAS: absorption spectroscopy; AI: artificial intelligence; ANN: artificial neural networks; ART: assisted reproductive technology; BPNN: back propagation neural network model; DT: decision tree; MLP: multilayer perceptron; PESA: percutaneous epididymal sperm aspiration; RBFN: radial basis function network; SRNN: simple recurrent neural network; SVM: support vector machines; TSE: testicular sperm extraction; WHO: World Health Organization

ARTICLE HISTORY

Received 14 October 2015
Revised 21 March 2016
Accepted 22 March 2016

KEYWORDS

Artificial neural networks;
biochemical markers; human
seminal plasma; prediction

Introduction

Seminal fluid is the secretion from many glands comprised of several organic and inorganic compounds including free amino acids, proteins, fructose, glucosidase, zinc, and other scavenging elements like Mg^{2+} , Ca^{2+} , K^+ , and Na . Mg is responsible for the modulation of erectile function and sperm motility in the vas deferens and the female reproductive tract. Zinc ions contributed by the prostate at the time of ejaculation aid in the production of semen coagulum which is an elastic substance that exerts a physical constraint on sperm motility. After liquification of semen in the female reproductive tract, zinc ions reinitiate sperm motility. Semen is 50% proteins by weight and the seminal plasma proteins are required for proper sperm function and its interaction inside the female genital tract. Some

peptides and proteins act to signal the female immune system for sperm rejection or tolerance and ultimately influence the relative intrinsic fertility of the couple. To develop novel approaches towards proper diagnosis to male infertility, an overall understanding of the biochemical and molecular composition and its role in regulating sperm quality is highly desirable. Laboratory procedures pertaining to the evaluation and estimation of the fertility associated biomolecules in semen are exhaustive and time consuming. For this purpose, major biochemical markers were elucidated and predicted with artificial neural network (ANN) models.

Three different ANN models like Multilayer Perceptron (MLP), Decision Tree (DT), and Support Vector Machines (SVM) have been used to predict

CONTACT T. B. Sridharan  tbsridharan@vit.ac.in  School of Biosciences and Technology, VIT University, Vellore, Tamil Nadu 632014, India.

Color versions of one or more of the figures in the article can be found online at www.tandfonline.com/iaan.

© 2016 Taylor & Francis

various parameters involved in the male fertility potential [Gil et al. 2012]. Furthermore, ANN has also been used to predict various medical applications. The major advantage of ANN in predicting the parameters for biological samples is its potential to sustain clinical decision using a cheap and sensitive tool while handling the maximum number of samples within the time limit. It is easy to compare the results with the existing sample database which makes it helpful in creating a biological and biochemical parameter dataset as a single software solution [Lisboa and Taktak, 2006].

A large number of classifiers are available within artificial intelligence (AI), including ANNs and SVMs [Polat et al. 2009; Conforti and Guido, 2010]. Patients with azoospermia male infertility are typically availed techniques like testicular sperm extraction (TSE), percutaneous epididymal sperm aspiration (PESA), that are used in assisted reproduction for enhancing the fertility [Devroey et al. 1994; Silber et al. 1996]. The success rate for assisted reproduction increases using these surgical techniques. The efficiency of many of these techniques varies in the presence of fibrosis and vascularisation, which can burden and hence make it prohibitive within the framework of assisted reproduction [Schlegel and Su 1997; Tournaye et al. 1997].

Prediction of various outcomes of assisted reproduction by using regression models is not an easy task, but the development of an accurate model is possible by the application of ANN following a Bayesian algorithm. The results of prediction techniques for the IVF outcomes and diagnosis of male infertility by ANNs is still not reliable [Corani et al. 2013]. Scientists have developed three different AI methods to predict semen quality assuming that the semen has been influenced by environmental and stress related issues [Corani et al. 2013]. Our previous report on the development of an ANN model by using a back propagation neural network model (BPNN) to predict the concentration of Zn in normospermia samples using the semen parameters as input was found to be useful in elucidation and validation of male infertility [Vickram et al. 2013]. We used six different input parameters and only one output (predicting the concentration of Zn) and finally

compared the prediction results with the original results obtained through atomic absorption spectroscopy (AAS); when the number of variables in the model increases, the ease and the accuracy of the prediction decreases [Bustillo et al. 1993; Jurisica et al. 1988]. The most noteworthy traits and their relationship in influencing the pregnancy success rate followed by IVF and diagnosing the male fertility potential are currently accepted as holding a breakthrough that challenges the experts in the AI field [Ruey-Shiang et al. 2011].

Results

Initially, 177 samples were analyzed for predicting four different outputs that included protein concentration, fructose concentration, glucosidase activity, and Zn concentration with semen parameters as inputs. The data for all the biochemical parameters were initially evaluated by manual methods. Semen parameters were elucidated for all the categories and tabulated (Table 1). Three different neural networks were used to predict four major outputs.

Back propagation neural network (BPNN)

Back-propagation networks were composed of layers of neurons. The input and output layer were connected. The BPNN consists of an input layer, one or two hidden layers, and an output layer. A schematic diagram of a BPNN with n input nodes, r output nodes and a single hidden layer of m nodes are shown in Figure 1. All the connections have multiplying weights associated with them.

Radical basis function network (RBFN)

RBFN is a three layer feed-forward network that contains one input layer, one middle layer, and one output layer as shown in the model (Figure 2). The first layer is the input layer for submission of data and it acts as the node for the connection to the network. The second layer is a singular hidden layer which is different from the

Table 1. Comparison of semen parameters with various categories of semen samples.

Semen category	Volume (ml)	pH	Sperm concentration (millions/ml)	Total motility (%)	Rapid progressive motility (%)	Normal morphology (%)
Oligoasthenospermia ($n=35$)	2.1±0.2	7.7±0.0	8.2±1.9	7.2±1.5	3.2±0.9	20.3±1.2
Asthenospermia ($n=35$)	2.8±0.5	7.7±0.0	29.4±4.7	8.2±1.9	2.6±0.5	20.1±3.9
Azoospermia ($n=22$)	1.9±0.5	7.7±0.1	NIL	NIL	NIL	NIL
Normospermia ($n=34$)	3.9±0.7	7.4±0.3	79.4±10.2	61.8±8.3	26.2±2.5	37.7±5.8
Oligospermia ($n=34$)	2.1±0.5	7.5±0.1	2.2±0.1	20.8±0.5	30.2±1.6	20.8±2.1
Control ($n=17$)	4.2±0.8	7.5±0.0	110.1±14.2	58.6±7.7	30.0±4.6	40.7±4.5

All the samples were represented by mean ± standard error of mean.

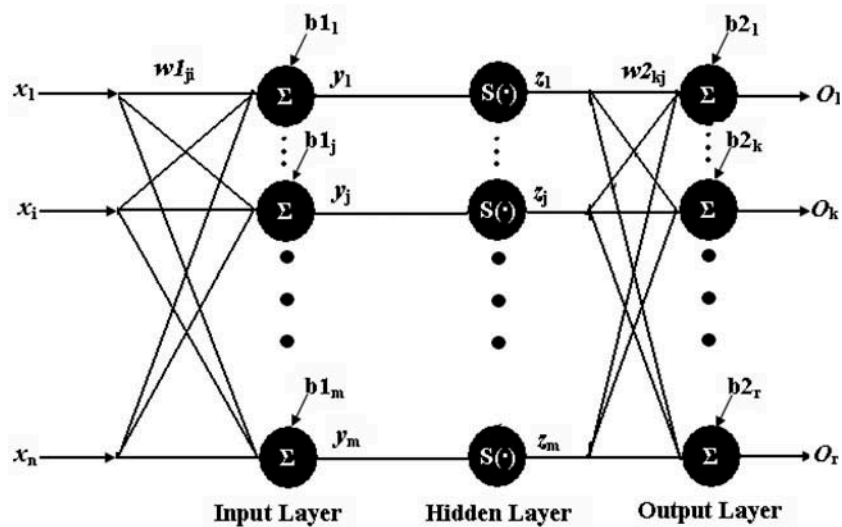


Figure 1. Back propagation neural network model (BPNN). Back propagation networks were composed of layers of neurons. The input and output layer were connected. The BPNN consists of an input layer, one or two hidden layers, and an output layer. All the connections have multiplying weights associated with them.

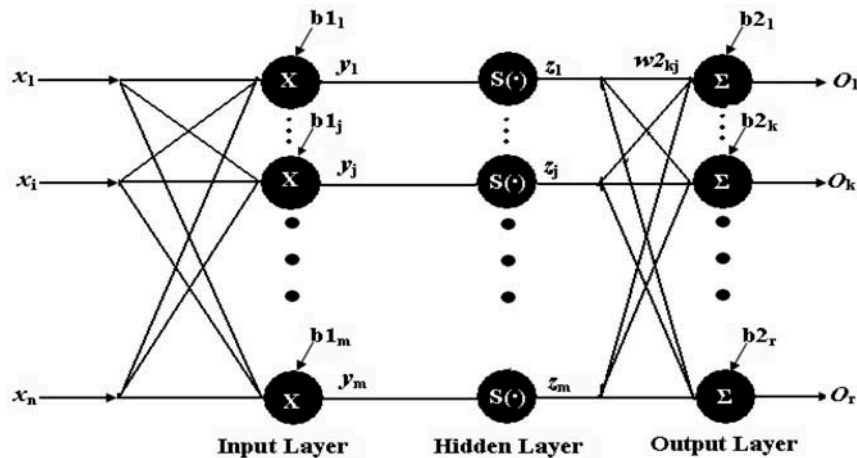


Figure 2. Radial neural basis function network model (RBFN). RBFN is a three layer feed-forward network that contains one input layer, one middle layer, and one output layer as shown in the model. The first layer is the input layer for submission of data and it acts as the node for the connection to the network. The second layer is a singular hidden layer which is different from the hidden layer of BPNN having multiple layers. The input nodes pass to incoming input vector to the hidden nodes. The connection between the hidden nodes and the input nodes is not weighed. But, the connection between hidden nodes and output nodes is weighed.

hidden layer of BPNN having multiple layers. The input nodes pass to incoming input vector to the hidden nodes. The connection between the hidden nodes and the input nodes is not weighed. But, the connection between hidden nodes and output nodes is weighed. The hidden layer acts as an input space for non-linear mapping to hidden space. The third layer is an output layer for linear mapping to obtain an output value. A radial function is real valued function whose value depends only on the distance from the origin. Their characteristic feature is that their response decreases (or increases) monotonically with distance from the central point.

Simple recurrent neural network (SRNN)

SRNN follows the same algorithm as the BPNN algorithm except that SRNN has an extra layer called the simple recurrent layer which has been shown in the model (Figure 3).

In this algorithm, first the input nodes were connected to the hidden layer. Again the hidden layer was connected to an extra layer called the context layer or recurrent layer, where it gained weight and got connected to the hidden layer through another route as clearly shown in the model (Figure 3).

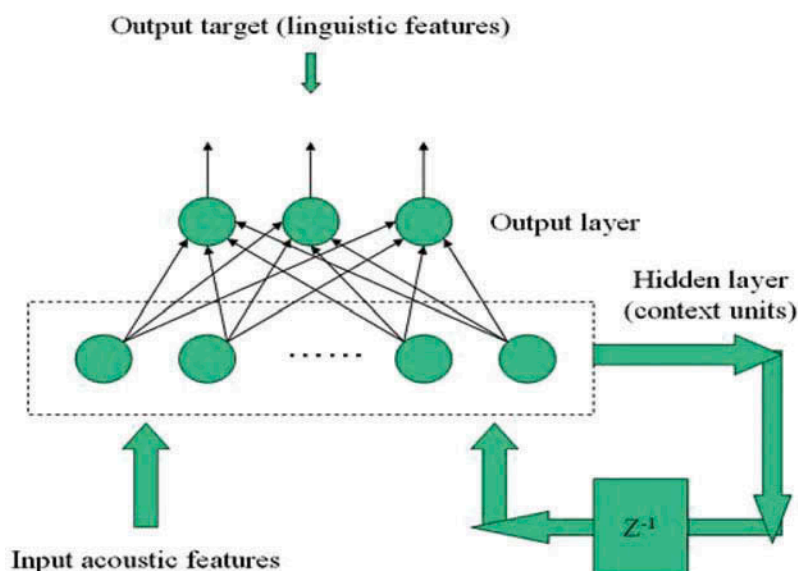


Figure 3. Simple recurrent neural network model (SRNN). SRNN follows the same algorithm as the BPNN algorithm except that SRNN has an extra layer called simple recurrent layer which has been shown in the model. In this algorithm, first the input nodes were connected to the hidden layer. Again the hidden layer was connected to an extra layer called context layer or recurrent layer, where it gained weight and got connected to the hidden layer through another route as clearly shown in the model.

A designed model of an ANN was used to predict four different outputs. The semen parameters including volume, pH, sperm concentration, total motility, progressive motility, and normal morphology were taken as inputs connected with the ANN model to predict the levels of fructose and zinc, glucosidase activity, and protein concentration.

For training the three different networks, 70 samples were taken. From each category, oligospermia, oligoasthenospermia, asthenospermia, azoospermia, normospermia and the control, ten samples were taken and used to train the network. The input and output for the 70 samples were given and the network was trained to predict the test. Later 71 samples were used to validate the model for prediction. Finally, 36 samples were used to predict the outputs for which only inputs were given to the models. The absolute error in prediction such as the minimum error and maximum error for prediction was tabulated then compared to the manual method (Table 2).

BPNN, RBFN, and SRNN models and outputs

The absolute error in predicting four different outputs using the BPNN model was acceptable as shown (Figure 4). The mean absolute error for the BPNN model was 0.025, -0.080, 0.166, and -0.057 for protein, fructose, glucosidase, and zinc, respectively. Zn prediction was found to be with very less error when compared with all the other parameters in the BPNN model.

The absolute error in predicting four different outputs by the RBFN model was acceptable as shown (Figure 4). The mean absolute error for the RBFN model was 0.961, -0.333, -0.039, and -0.110 for protein, fructose, glucosidase, and zinc, respectively. This error was not equal to the required standard when compared to the BPNN model prediction. The error for predicting the protein concentration was found to be very high with RBFN when compared to the BPNN model.

Table 2. Comparison of error of neural network models for different outputs.

Absolute error for different ANN models	Total semen protein concentration (mg/ml)		Fructose concentration μ mol/ejaculate		Glucosidase activity (U/ml)		Zn concentration (mg/ml)	
	Min Error	Max Error	Min Error	Max Error	Min Error	Max Error	Min Error	Max Error
BPNN model	-6.600	4.00	-2.000	4.000	-2.000	2.500	-0.3800	0.0800
RBFN model	-6.000	7.60	-4.800	6.000	-5.000	6.000	-0.6800	0.0790
SRNN model	-6.600	7.00	-6.000	7.000	-6.700	4.000	-0.9770	0.8200

ANN: artificial neural network; BPNN: back propagation neural network; RBFN: radical basis function network; SRNN: simple recurrent neural network.

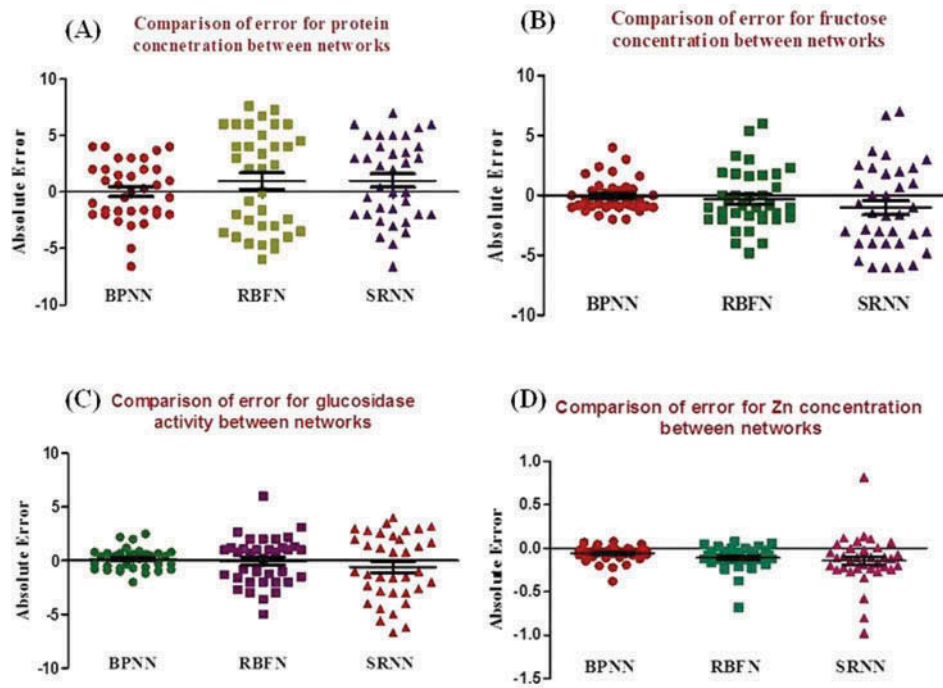


Figure 4. Comparison of error in prediction for various output parameters. (A) The p values were 0.06848, <0.0001, and 0.0746 for back propagation neural network model (BPNN), radial neural basis function network model (RBFN), and simple recurrent neural network model (SRNN), respectively, for predicting protein concentration. RBFN showed significant difference, whereas BPNN and SRNN showed no significant difference. (B) The absolute error was calculated between the manual examination and predicted values with three different networks for fructose concentration. All three tests showed that BPNN was significantly different, in comparison to RBFN and SRNN that showed no significant difference. (C) The absolute error was calculated between the manual examination and predicted values with three different networks for glucosidase activity. The p values were 0.5947, 0.7260, and 0.0671 for BPNN, RBFN, and SRNN, respectively. (D) The absolute error between the manual examination and the predicted values with three different networks were calculated for Zn concentration. The p values were <0.0001, <0.0001, and 0.0052 for BPNN, RBFN, and SRNN, respectively. All networks showed a significant difference when compared.

The absolute error in predicting four different outputs by the SRNN model was also acceptable as shown (Figure 4). The mean absolute error for the SRNN model was 1.017, -1.033, -0.589, and -0.140 for protein, fructose, glucosidase, and zinc, respectively. The absolute error for the SRNN model was higher than the BPNN model. The prediction pattern for protein concentration was considerably increased for RBFN and SRNN when compared to the BPNN model. This likely reflects the larger deviation in the parameter values.

The absolute error was calculated between the manual examination and predicted values with three different networks for whole protein concentration. D'Agostino and Pearson omnibus normality test was used to compare the p values between the three networks. The p values were 0.06848, <0.0001, and 0.0746 for BPNN, RBFN, and SRNN, respectively. RBFN showed significant difference, whereas BPNN and SRNN showed no significant difference. The mean absolute error for BPNN was 0.0250 which was much

less when compared to the other networks (Figure 4A). The absolute error was calculated between the manual examination and predicted values with three different networks for fructose concentration. D'Agostino & Pearson omnibus normality test, KS normality test, and Shapiro-Wilk normality test were done to compare the p values between the three networks. All three tests showed that BPNN was significantly different, in comparison to RBFN and SRNN that showed no significant difference. The mean absolute error for BPNN was -0.08056 which was much less when compared to other two networks (Figure 4B).

The absolute error was calculated between the manual examination and predicted values with three different networks for glucosidase activity. D'Agostino & Pearson omnibus normality test was carried out to compare the p values between the three networks. The p values were 0.5947, 0.7260, and 0.0671 for BPNN, RBFN, and SRNN, respectively. All networks showed no significant difference when compared to each other.

The mean absolute error for BPNN was 0.1661 which was much less than that of the other networks (Figure 4C).

The absolute error between the manual examination and the predicted values with three different networks were calculated for Zn concentration. D'Agostino & Pearson omnibus normality test was done to compare the p values between three networks. The p values were <0.0001 , <0.0001 , and 0.0052 for BPNN, RBFN, and SRNN, respectively. All networks showed a significant difference when compared. The mean absolute error for BPNN was -0.05782 and was much less than other two networks (Figure 4D).

Discussion

ANNs with their pattern recognition and modelling capabilities have been successfully used in the field of biomedicine to help in the diagnosis of hepato-biliary disorders [Hayashi et al. 2010], coronary disorders [Azuaje et al. 1999], estimation of lead concentration in grasses [Dimopoulos et al. 1999], and prediction of membrane fouling during nanofiltration of ground and surface water [Shetty and Chellam 2003]. Predicting assisted reproduction outcome is not easy by regression models, but ANNs can be used in developing a more precise model with the application of Bayesian algorithm. The results of prediction techniques for the IVF outcomes and diagnosis of male infertility by ANNs is still not reliable [Corani et al. 2013]. Gil and researchers [2012] developed three different AI approaches to predict the semen quality assuming that the semen has been influenced by various environmental and stress related factors. Their application should prove useful in elucidating and validating the diagnosis of male infertility [Gil et al. 2012].

Recently, Gil et al. [2012] evaluated the performance of various artificial networks and their application in the prediction of male fertility potential, aimed at checking the semen quality, and also used non linear statistical techniques that may allow a better approach to address the complexity of the problem. However, to our knowledge, this is the first report on the use of ANN for the accurate prediction of Zn concentration in fertile human semen samples. The ANN model is mainly used to predict the quality of spermatozoa before the diagnosis of male infertility and to compare the results with a standard regression model. Authors concluded that this model's prediction is of lower error than any other standard models [Murat and Dogan 2004]. In our research, we mainly focussed on designing three different neural networks like BPNN, RBFN, and SRNN to predict four

different outputs like protein content, fructose concentration, glucosidase activity, and zinc concentration.

The amount and the number of parameters used to predict semen quality are important factors in the application of AI for improving the accuracy and to elucidate the most influential parameter [Gil et al. 2009; Gil and Johnsson 2010a; 2010b; Gil et al. 2011; Subashini et al. 2009]. Various network algorithms were adopted to predict the quality of the sperm as a function of the modern lifestyle. The deterioration of major semen parameters like sperm concentration, total motility, and normal morphology are correlated with modern lifestyle factors [Yatsenko et al. 2012]. In our study, we used three neural networks BPNN, RBFN, and SRNN to predict protein concentration, fructose concentration, glucosidase activity, and zinc concentration with the help of semen analysis parameters as inputs (Figure 4). Available literature reveals that AI can help in the prediction of sperm quality including sperm morphology and sperm kinetics. This can be aided by incorporating various seminal parameters as inputs and biochemical properties as outputs.

Conclusion

Seminal biochemical markers can be treated as a proxy measure of male fertility and they have long been measured biochemically employing various analytical techniques. These add a little more on the expenditure list of AI procedures. The newly developed BPNN can be used to predict various biochemical markers with semen parameters as inputs. Greater accuracy achieved with BPNN models supports its potential usage in fertility centres.

Materials and methods

Semen sample collection

Semen samples were collected from the Andrology Department, Bangalore Assisted Conception Centre Private Ltd. A total of 177 semen samples were collected for this research (both fertile and infertile samples) and immediately processed to prepare a semen analysis report, based on the protocol of the World Health Organization [WHO 2010]. Based on the report, semen samples were categorized into oligoasthenospermia ($n=35$), asthenospermia ($n=35$), azoospermia ($n=22$), normospermia ($n=34$), oligospermia ($n=34$), and control ($n=17$).

Semen sample inclusion criteria

Males leading normal lives with regular unprotected sex without conception for one year or more were included in the study.

Semen sample exclusion criteria

A brief medical history of the patients was performed before semen analysis. The patients who were already using supplementary antioxidants or any other medication for male infertility were not included in this study. The patients who came for the first time for this reason were accepted for this study. In addition, subjects with testicular varicocele, genital infection, leukocytospermia, sexually transmitted diseases, chronic illness and serious systemic diseases, alcoholism, or smoking history were excluded from the study because of their well-known high seminal reactive oxygen species levels that decrease antioxidant activity, which results in decreased motility and abnormal morphology.

Research ethics

This research study is part of a major research project, for which the human ethical approval and clearance was obtained from the VIT University Institutional Human Ethical Committee, Ref. No. VIT/UHEC-3/NO.11. We collected the samples from patients who were receiving

semen analysis at the BACC in Bangalore, Karnataka, India. These patients gave their verbal consent to participate in the study. We are not able to get written statement from the patients, and this was also approved by the VIT University Institutional Human Ethical Committee, Ref. No. VIT/UHEC-3/NO.11. The patients were not willing to give written information because male infertility in India creates issues personally as well as socially. They were not willing to reveal that they are infertile to anyone outside the centre. Upon documenting the verbal consent, we noted the sample donor's name, address, and background. To record their voice (they are agreeing to do work with their samples) we used a MP3 recorder. We have a unique ID for each patient, and using that, we recorded their consent without their name (using unique ID number).

Assay and analysis protocols

The major biochemical markers like total protein, fructose, glucosidase, and Zn were elucidated for proper diagnosis of male infertility. All the samples were centrifuged at 3,000 x g to remove spermatozoa and frozen at -20°C until used. Total protein in the semen was estimated using the standard Lowry method. Total fructose concentration and neutral glucosidase in seminal plasma were elucidated by the standard WHO

[2010] protocol. Zn concentration in human seminal plasma was measured by Atomic Absorption Spectroscopy (AAS). Then all the biochemical parameters were predicted with three different ANN models, BPNN [Gil et al. 2012], RBFN [Gil et al. 2012], and SRNN [Subashini et al. 2009].

Acknowledgments

The authors were very much thankful to VIT University management, Vice Presidents, Vice Chancellor, and BACC Healthcare management for their support to complete this work successfully.

Declaration of interest

There is no conflict of interest for this research.

Notes on contributors

Performed the experiments: ASV, RP, KA; Designed the experiments: KAR, MRP; Artificial neural networks work: RD; Scientific ideas, part of design, and corresponding author: TBS.

References

- Azuaje, F., Dubitzky, W., Lopes, L., Black, N., Adamson, K., Wu, X. and White, J.A. (1999) Predicting coronary disease risk based on short-term RR interval measurements: a neural network approach. *Artif Intell Med* 15(3):275.
- Bustillo, M., Stern, J.J., King, A. and Coulam, C. (1993) Serum progesterone and estradiol concentrations in the early diagnosis of ectopic pregnancy after in vitro fertilization-embryo transfer. *Fertil Steril* 59(3):668–670.
- Conforti, D. and Guido, R. (2010) Kernel based support vector machine via semidefinite programming: Application to medical diagnosis. *Computers & Operations Research* 37(8):1389–1394.
- Corani, G., Magli, C., Giusti, A., Gianaroli, L. and Gambardella, L.M. (2013) A Bayesian network model for predicting pregnancy after in vitro fertilization. *Comput Biol Med* 43:1783–1792.
- Devroey, P., Liu, J., Nagy, Z., Tournaye, H., Silber, S.J. and Van Steirteghem, A.C. (1994) Normal fertilization of human oocytes after testicular sperm extraction and intracytoplasmic sperm injection. *Fertil Steril* 62:639–641.
- Dimopoulos, I., Chronopoulos-Serelia, J.A. and Lek, S. (1999) Neural network models to study relationships between lead concentration in grasses and permanent urban descriptors in Athens city (Greece). *Ecol Model* 120:157–165.
- Gil, D. and Johnsson, M. (2010a) Supervised som based architecture versus multilayer perceptron and rbfn networks. *Expert Systems with Applications* 37(6):15–24.
- Gil, D. and Johnsson, M. (2010b) Using support vector machines in diagnoses of urological dysfunctions. *Expert Systems with Applications* 37(6):4713–4718.

- Gil, D., Johnsson, M., Chamizoj, J.M., Paya, A.S. and Fernandez, D.R. (2011) Review article: Modelling of urological dysfunctions with neurological etiology by means of their centres involved. *Applied Soft Computing* **11** (8):4448–4457.
- Gil, D., Johnsson, M., Garcia Chamizo, C., Paya, J.M. and Fernandez, D.R. (2009) Application of artificial neural networks in the diagnosis of urological dysfunctions. *Expert Systems with Applications* **36**(3):5754–5760.
- Gil, D., Jose Luis, G., Joaquin, J., Jose Gomez, M., and Magnus, J. (2012) Predicting seminal quality with artificial intelligence methods. *Expert Systems with Applications* **39**:12564–12573.
- Hayashi, R., Setiono and Yoshida, K. (2010) A comparison between two neural network rule extraction techniques for the diagnosis of hepatobiliary disorders. *Artif Intell Med* **20**(3):205–216.
- Jurisica, I., Mylopoulos, J., Glasgow, J., Shapiro, H. and Casper, R.F. (1988) Case-based reasoning in IVF: Prediction and knowledge mining. *Artif Intell Med* **12**(1):1–24.
- Lisboa, P.J. and Taktak, A.F.G. (2006) The use of artificial neural networks in decision support in cancer: A systematic review. *Neural Networks* **19**(4):408–415.
- Murat, G. and Dogan, S. (2004) A neural network approach for early cost estimation of structural systems of buildings. *International Journal of Project Management* **22**(7):595–602.
- Polat, K., Gunes, S. and Salih, R. (2009) A novel hybrid intelligent method based on c4.5 decision tree classifier and one-against-all approach for multi-class classification problems. *Expert Systems with Applications* **36**(2):1587–1592.
- Ruey-Shiang, G., Tsung-Chieh, J. and Shao-Ping, W. (2011) Integrating genetic algorithm and decision tree learning for assistance in predicting in vitro fertilization outcomes. *Expert Systems with Applications* **38**:4437–4449.
- Schlegel, P.N. and Su L.M. (1997) Physiological consequences of testicular sperm extraction. *Hum Reprod* **12**:1688–1692.
- Shetty, G.R. and Chellam, S. (2003) Predicting membrane fouling during municipal drinking water nanofiltration using artificial neural networks. *J Membrane Sci* **21**(7):69–86.
- Silber, S.J., van Steirteghem, A., Nagy, Z., Liu, J., Tournaye, H. and Devroey, P. (1996) Normal pregnancies resulting from testicular sperm extraction and intracytoplasmic sperm injection for azoospermia due to maturation arrest. *Fertil Steril* **66**:110–117.
- Subashini, T.S., Ramalingam, V. and Palanivel, S. (2009) Breast mass classification based on cytological patterns using rbfn and svm. *Expert Systems with Applications* **36**(3):5284–5290.
- Tournaye, H., Verheyen, G. and Nagy, Z.P. (1997) Are there any predictive factors for successful testicular sperm recovery? *Hum Reprod* **12**:80–86.
- Vickram, A.S., Raja, Das., Srinivas, M.S., Kamini, A., Jayaraman, G. and Sridharan, T.B. (2013) Prediction of Zn concentration in human seminal plasma of Normospermia samples by Artificial Neural Networks (ANN). *J Assist Reprod Genet* **30**:453–459.
- WHO (2010) WHO laboratory manual for the examination and processing of human semen, 5th ed World Health Organization. pp. 7–36,56–102.
- Yatsenko, A.N., O’Neil, D.S., Roy, A., Arias-Mendoza, P.A., Chen, R., Murthy, L.G., et al. (2012) Association of mutations in the zona pellucida binding protein 1 (ZPBP1) gene with abnormal sperm head morphology in infertile men. *Mol Hum Reprod* **18**(1):14–21.