Title: SEMEN PARAMETERS CAN BE PREDICTED FROM ENVIRONMENTAL FACTORS AND LIFESTYLE USING ARTIFICIAL INTELLIGENCE METHODS
Short Title: Predicting Semen Quality with Neural Networks

Jose L. Girela[1,*,†], David Gil[2,†], Magnus Johnsson[3], María José Gomez-Torres[1], and Joaquín De Juan[1].
[1]Department of Biotechnology, University of Alicante, Alicante, Spain
[2]Lucentia Research Group, Department of Computer Technology, University of Alicante, Alicante, Spain
[3]Department of Cognitive Science, University of Lund, Lund, Sweden

This study was partially funded by Vicerrectorado de Investigación, University of Alicante, Alicante, Spain (Vigrob-137).

* Correspondence: E-mail: girela@ua.es, jdj@ua.es.
[†]Both authors contributed equally to this work.

ABSTRACT
Fertility rates have dramatically decreased in the last two decades, especially in men. It has been described that environmental factors as well as life habits may affect semen quality. In this paper we use artificial intelligence techniques in order to predict semen characteristics from environmental factors, life habits and health status, as a possible Decision Support System that can help in the study of the male fertility potential. One hundred twenty three young healthy volunteers provide a semen sample that was analyzed according to the World Health Organization 2010 criteria. They also were asked to fulfill a validated questionnaire about life habits and health status. Sperm concentration and percentage of motile sperm were related to socio-demographic data, environmental factors, health status, and life habits, to determine the predictive accuracy of a Multilayer Perceptron Network, a type of Artificial Neural Network. In conclusion, we have developed an Artificial Neural Network that can predict the results of the semen analysis, based on the data collected by the questionnaire. The semen parameter that is best predicted using this methodology is the sperm concentration. Although the accuracy of motility is slightly lower than concentration, it is possible to predict it with a significant accuracy. This methodology can be a useful tool in order to early diagnosis of patients with seminal disorders or in the selection of candidates to become semen donors.

Summary sentence: Artificial Neural Networks predicts, with high classification accuracy, seminal parameters of young and healthy volunteers using the health status, lifestyle and exposure to environmental factors as inputs variables.

Keywords: Semen quality, life habits, Supervised Learning, Artificial Neural Network, Decision Support System.

INTRODUCTION
In the last two decades there has been a notable decline in fertility rates [1-3]. It was considered that this decline is due to changes in behaviour related to economic aspects, with the incorporation of women into labour and the consequent delay in the age at which you decide to

have offspring, and the widespread use of contraceptives [4-6]. Although it is clear that the social aspect has contributed significantly to this global decline in fertility some authors suggest that occurred synchronously with the deterioration of reproductive health caused by adverse biological factors [7].

Rates of demand for assisted reproduction treatment show an increase of situations where the male factor is altered [8, 9]. Over the past decades since the publication of a meta-analysis directed by Elisabeth Carlsen [10], remains a debate about the possibility of a decline in seminal quality. Numerous studies show a decrease in semen parameters of men [11-14], although there are studies that found no evidence of that decline [15, 16]. About the causes of this possible decline in semen quality that affect male fertility, several factors have been considered, from an increase in the incidence of male reproductive diseases [17, 18], environmental or occupational factors [19, 20], to a certain lifestyle [21, 22].

Semen analysis is the cornerstone of the male study [23]. Although semen analysis alone cannot determine whether a male can have offspring, it is a good predictor of male fertility potential [24-27]. Semen analysis is also necessary to evaluate candidates to become semen donors [28-31].

In this paper, we study the male fertility, approaching the problem from the perspective of possible influence of environmental factors and life habits in semen quality. To do that we use Artificial Intelligence (AI) techniques to produce a Decision Support Systems (DSS) that can help in the prediction of semen parameters [Figure 1].

In the last two decades, the use of AI has also become widely accepted in medical applications. Many of those applications have advanced to the materialization of Expert Systems and DSSs in several different areas. In Andrology, there have been only a few approximations [32-36]. Although these works show the potential of Artificial Neural Networks (ANN), we believe that they have not been enough developed. Therefore, we found necessary to evaluate this methodology applied to this field.

Applications built with ANN architecture include several advantages, such as generalization, facility in the optimization (which makes it more flexible for non-linear modeling of large data sets) and accuracy for predictive inference with potential to DSS (this is the reason for using this architecture). Moreover, these models can make knowledge dissemination easier by providing an explanation [37]. ANNs have been used to improve the classification tasks because of its property called black-box learning being the most popular methods for classification problems [38, 39] with better performance over other methods [40, 41].

AI methods have been commonly used to improve the accuracy. This is probably also one of most popular examples of AI uses. However, these architectures are also very suitable to deal with big data, as well as heterogeneous information.
In the evaluation of the infertility patient may be useful to apply AI methods, not only to improve the accuracy, but also to select the best features since the number of variables is huge as well as the amount of information used [42-46].

Knowledge discovery in databases, data mining or extractions of patterns from large data sets (or recently summarized as big data) will be the natural next step in this research area.

The objective of this paper is to develop a DSS based on Artificial Intelligence methods that can predict the results of the semen analysis from the answers of a questionnaire (see Supplemental Data, available online at www.biolreprod.org), and be a useful tool in order to early diagnosis and cribbage of patients with seminal disorders or in the selection of candidates to become semen donors.

MATERIALS AND METHODS

Study population

To perform the present study we recruit volunteers among the students of the University of Alicante, as a population of young and healthy males. It has been previously established that the use of volunteers does not introduce selection bias in male fertility studies [47], and as we were interested in the impact of environmental factors and lifestyle in semen parameters we decided not to study individuals with known or suspected reproductive disorders (i.e. patients of fertility clinic). Moreover, it is also important to highlight that the study population of this work presents close characteristics to regular candidates to become semen donors (i.e. young male, university students) [48]. 123 volunteers' between 18 and 36 years old participated in the study, but only data from 100 individuals were used due to incomplete information of some of the subjects.

Semen analysis

After being informed, volunteers were asked to provide a semen sample after 3 to 6 days of sexual abstinence, and a semen analysis according to WHO guidance was done [49]. All semen samples were processed by the same person with the same equipment, and analysis of motility and concentration was realized within the first 60 minutes after semen collection. All samples were obtained and processed with the approval of the Institutional Review Board of the University of Alicante.

In this study we use the seminal analysis results as a control to evaluate the accuracy in the prediction of the diagnosis of the ANN developed. These ANN will be only based in the answers to the questionnaire of the individuals who participate.

Variables of the questionnaire

On the day of the analysis, the volunteers were asked to fulfill a questionnaire about life habits and health status. They were asked about 34 items (table 1).

After a pre-processing using Decision Trees (DT), which is another Artificial intelligence method, we found that not all of the 34 questions in the questionnaire developed were relevant and then necessary in the classification process [50]. In table 2 are shown those questions that were found relevant for each one of the seminal parameters studied. Only the age was present in both sets of questions. In this regard, the seminal alterations studied in this paper which are sperm concentration and motility, are produced by different factors and in different moments. The sperm concentration is determined by the efficiency of the germinal epithelium while motility reflects the maturation of the spermatozoa produced. Therefore we decided to develop

two specific neural networks, one of them to predict the sperm concentration (ANN1) and the other one to predict the sperm motility (ANN2) in order to obtain the most accurate results.

Table 2 also shows a description and the range of values of the variables used in the study as well as the values normalized after the treatment we have done to the different database input fields. We have converted the input data into a range of normalization according to the follow rules:
a) Numerical variables such as age are normalized onto the interval (0 - 1).
b) The variables with only two independent attributes are prearranged with binary values (0, 1).
c) The variables with three independent attributes, such as "Vaccines received", "High fevers in the last year" and "Smoking habit" are prearranged using the ternary values (-1, 0, 1).
d) The variables with four independent attributes, such as "Season in which the analysis was performed" or "Marital status" are prearranged using the four different and equal distance values (-1, -0.33, 0.33, 1).

ANN architecture

An ANN is a mathematical model inspired by biological neural networks. A neural network consists of an interconnected group of artificial neurons, and it processes information using a connectionist approach to computation. In this context, a neuron is the basic computation unit. In these units a series of mathematical operations are developed. Afterwards they decide the next step in the computation pathway depending on the results obtained, which is called the activation function.

In this study we have used a Multilayer Perceptron (MLP) network [38, 51, 52] which as a special sort of ANN comprises an architecture of several layers of neurons (therefore, it keeps the resemblance with the biological brain since it is a bioinspired method).

In our case, which is very common, we chose three layers (where each one is fully connected to the next one): an input layer that receives external inputs, one hidden layer, and an output layer which, normally, and it also happens in our case, it generates the classification results [see Figure 2].

After the first level of neurons (in the input layer) the rest of neurons form computational elements with a nonlinear activation function. In order to summarize the MLP functioning, the strategy of the network is that when data are presented in the input layer, the remainder neurons run calculations in the consecutive layers until an output value is achieved for the output neurons which will specify the correct class for the input data.

Figure 2 indicates how the hidden layers neurons compute weighted sums of their inputs and add a threshold. The resulting sums are used to calculate the activity of the neurons by applying a sigmoid activation function.

This process is defined as follows:

$$v_j = \sum_{i=1}^{p} \omega_{ji} x_i + \theta_j \ , \ \ y_j = f_j(v_j)$$

where υ_j is the linear combination of inputs $x_1, x_2, \ldots, x_p$, and the threshold θ_j, ω_{ji} is the connection weight between the input x_i and the neuron j, and f_j is the activation function of the j_{th} neuron, and y_j is the output. The sigmoid function is a common choice of activation function. It is defined as:

$$f(t) = \frac{1}{1+e^{-t}}$$

A single neuron in the MLP is able to linearly separate its input space into two subspaces by a hyperplane defined by the weights and the threshold. The weights define the direction of this hyperplane whereas the threshold term θ_j offsets it from origo.
This is essentially what it does MLP very suitable for the classification problems.
The MLP network uses the backpropagation algorithm [53], which is the most suitable algorithm for similar works [44, 50].

The different layers of the architecture MLP is assembled as follows:
a) Layer 1: It is built automatically from the input vector, which in our case are of the answers of the questionnaire shown in table 2.
b) Layer 2: To decide the number of hidden neurons for this layer is the hardest issue in the network's architecture. This number represents the equilibrium between a good accuracy and the possibility of over fitting. In the matter of fact, the precise number of neurons in a hidden layer will improve the capacity of the network of generalization from new data notably [54].
c) Layer 3 is the output (Classification) layer. The output corresponds to the different problem classes. In our example there are two outputs, normal or altered.

The weights and the threshold of the MLP are calculated during an adaptation process. In the results section the number of hidden neurons of the MLP will be established.

All the computations of the neural network was developed using the WEKA software [55]. Weka includes most of the machine learning methods and it is Open Source software developed by the University of Waikato in New Zealand.

ANN experimentation

Since the amount of data to examine our system is normally not very big, a method called cross validation is adapted. The idea is to divide the entire set of data into two subsets, namely training set and test set. The first one, training set, is used to determine the system parameters (also to configure the hidden layer explained in the former section). The test set is used to assess the prediction accuracy and the network generalization. This method has been used very often to measure the generalization of neural networks.

The method is called n-fold cross validation since this process is performed n times. For every test, the test subset is chosen randomly from the initial pull and the remaining form the training data.

To evaluate the performance we obtain some measures as classification accuracy, sensitivity, specificity, positive predictive value, negative predictive value and a confusion matrix. A

confusion matrix [56] contains information about actual and predicted classifications done by a classification system.

RESULTS

Baseline semen characteristic and laboratory results

Table 3 shows the mean of the seminal parameters from the study population. Of all men surveyed, 38% had some alteration in sperm parameters. By studying in detail the individuals with altered parameters the most frequently found seminal alteration was asthenozoospermia in 18% of cases, followed by oligozoospermia in 8%. The combined alterations astenoteratozoospermia and oligoasthenozoospermia have shown a frequency of 4% in both cases. 3% of individuals showed oligoasthenoteratozoospermia. Finally, the less frequent alteration was teratozoospermia without other alterations, present only in one individual.

ANN performance

In order to keep a good generalization and consequently a high accuracy, without getting over fitting, the experiments carried out shown that a low number of neurons for the Layer 2 (the hidden layer) lead to a poor performance for both training and test sets. In the opposite edge, a high number of neurons perform very well for training and test sets, although the risk of over fitting is high. The compromise between these two options leads us to find the optimal solution for this layer in 6 neurons for each one of the neural networks developed.
The output (Classification) layer has two classes for classification/prediction: normal and altered based on WHO reference values [49]. As we explained regarding to table 2 we have decided to develop two Neural Networks in order to analyze the measure of the seminal parameters studied: sperm concentration (ANN1) and motility (ANN2) since both nets used different variables except the age which is common.

To assess the generalization of the network we have applied a ten-fold cross-validation method for the performance assessment of every network. The data have been divided in ten sets (S1, S2, ..., S10) in order to carry out ten experiments. This validation was optimal and stabilized after the first 1000 epochs.

Table 4 shows the confusion matrix provided by our classifiers for each of the seminal parameters. With the results of confusion matrix it is acquired the following parameters of the ANN performance: Classification accuracy, sensitivity, specificity, positive predictive value and negative predictive value. Table 5 shows the values obtained for each seminal parameter.
ROC curves [Figure 3] for each of the seminal parameters studied were made to show the accuracy in the prediction of the semen characteristic by the ANN using only as inputs the variables in the questionnaire. Sperm concentration is the seminal parameter that shows the highest prediction accuracy. Although the accuracy of motility is slightly lower than concentration, it is possible to predict it with a significant accuracy.

DISCUSSION

In this paper, we have presented the functioning of an artificial neural network, specifically an MLP as a new approach in the prediction of the male fertility potential from life habits and environmental factors. The experiments have been carried out with the data obtained from 100

volunteers' between 18 and 36 years and we show the relationship of those factors with semen parameters.

The experiments carried out have achieved very good accuracy parameters for MLP. Among them, classification accuracy, sensitivity, negative predictive value obtain the highest precision values whereas specificity and positive predictive value present lower percentages. This situation occurs even when the classification and the confusion matrix are good. We have to find the explanation in the input data since the study population shows an imbalanced distribution [57] because there are more individuals with normal semen parameters than the ones with alterations. This is very common in AI applied to classification in general and in medicine in particular. The development of this system and its gradual incorporation to medical centers will improve those values.

Although there are other good classifiers in the AI field, such as decision trees [43], support vector machines [58] or even hybrid methods that combine artificial neural network (ANN) and fuzzy neural network (FNN) [59], the advantage of MLP is the power of its generalization and extrapolation to many different areas. This is especially interesting when you deal with huge and heterogeneous amount of information, and you must evaluate it in the most objective manner possible. In addition, MLP has been used in many other fields [39, 40] with a high accuracy.

Therefore, it shows a very interesting approach for the prediction of the male fertility potential. In the field of reproductive biology and in particular in male factor studies, ANN has been used to predict the results of IVF/ICSI [32, 33, 35], to assess sperm morphology [60] and to predict the presence of sperm in testes of men with nonobstructive azoospermia [36, 61].

Neural network seems to be an alternative to more expensive laboratory tests, at least during the initial moment of the fertility study of the couple, as well as in the selection of semen donors. It is also important to highlight that the study population of this work presents close characteristics to regular candidates to become semen donors (i.e. young male, university students) [48]. It can also help in the prioritization and selection of the next steps of the infertility treatment, focusing or not on the male partner and avoiding painful and expensive examinations on the female. The authors suggest incorporating this tool in the protocols of the diary andrological evaluation, in order to obtain a first estimate of the seminal profile.
Other studies in different areas, especially in the field of medicine, have shown several lacks and limitations. Most of those problems are related to the complexity of the information treated due to its diversity and heterogeneous data types. In our opinion, the problem is that previous works [62] have approached the study of the relationship between environmental factors and lifestyle with the seminal quality using linear statistical techniques. In this paper, we develop the use of non-linear techniques that may allow a better approach of the complexity of the problem [63, 64].

In relation with the different factors used to develop the two neural networks, only age was present in both of them, showing specificity between these particular variables and the semen parameter they affect. In ANN1, that predicts the sperm concentration, one of the factors used is the season in which the analysis was performed. This seems to be in accordance with previous studies that show differences in the sperm concentration and the total sperm count between seasons [65, 66]. Another factor present in this neural network is the presence of childish diseases

(i.e. mumps, chicken pox, polio). These diseases are usually accompanied by high fever and in some cases like mumps are related to other reproductive problems like orchitis, factors that are known to impair the sperm production in the testis [19, 67]. In relation with this factor febrile illness is present in this neural network, and it is also supported by previous studies that have shown the relation with the sperm concentration [68]. Another factor related with heat stress in the testis, that is present in this neural network, is time spent sitting and it has been previously reported to increase the risk of oligozoospermia [69]. Smoking habit and alcohol consumption information are factors used by the neural network in the classification process. The relationship of tobacco with sperm concentration is shown in several studies [70, 71]. Meanwhile, it seems to be difficult to relate alcohol consumption with an impairment of sperm counts as a single factor [72]. There are two factors, history of surgical intervention and accident or serious traumas (but non testicular trauma), without previous studies of their impact in sperm production. Therefore, we suggest further research that can show the possible mechanism of their relation with sperm concentration. In the second neural network ANN2, the one that predicts the sperm motility, we found Body Mass Index (BMI) as a first factor. Although previous studies show this relationship [73] there are contradictory results, and further studies are needed to evaluate the impact of BMI in semen quality [74, 75]. Another contradictory factor is the marital status. Although it is present in most of the studies of environmental factors, there are no evidences of its influence over semen quality and in particular over sperm motility [76]. Radiation exposure (i.e. medical x-ray) is known to have a potential detrimental effect in testis function [77] and is included as an input variable in this network. Suffering of allergy is also included as a factor in the neural network. It has been previously suggested that this factor may affect male fertility potential but the mechanism of action is not completely known. Another factor included, with contradictory results, is the use of hot baths [78]. For the other factors in this network there is no scientific evidence of their influence in semen quality, suggesting the necessity of further investigation. Among them there is the number of sibling, the exposure to vaccines, and the average of hours sleep by day.

The direct application of AI techniques in previous studies by the authors [44-46, 50] has made possible to adjust and to configure the architectures of ANNs to satisfy the needs of the data analyzed in this study. To allow other investigators to access to the artificial neural networks developed, we offer an open access software implementation through the website url <http://dbt.ua.es/en/research/vida.html>**.**

The next step is to solve the effect of imbalanced classes on classification performance. It has been pointed out that this is one of the drawbacks in the development of MLP, especially in the medicine area [79]. The increase in the data collection, repositories such as data warehouses, will lead to the use of *machine learning* methods, as well as *data mining*, which may provide a more precise correlation between seminal data and the information collected by means of the questionnaires. In accordance with this, we have donated the whole database of this study to the most important repository of databases used for machine learning tasks, UCI machine learning (http://archive.ics.uci.edu/ml/), at University of California, Irvine. The purpose of it is to allow other researchers to access this information without limitations.

In conclusion, to our knowledge, this is the first time that MLP has been used to evaluate the relationship between life habits and semen quality. The neural networks developed in this study

show very high prediction accuracies, being slightly superior in the case of sperm concentration prediction. The construction of two architectures different for the prediction of the two seminal parameters studied, sperm concentration and motility, has help us to identify those factors that affect specifically to each of the two seminal parameters. In this regard, further studies should be carried out in order to propose additional recommendations to improve the male fertility potential

ACKNOWLEDGMENTS
The authors would like to acknowledge Daniel Cortes Blasco, from *i-think.es*, for his assistance in the development of the software implementation to provide access to the Artificial Neural Network described in this study.

REFERENCES
1. Inhorn MC. Global infertility and the globalization of new reproductive technologies: illustrations from Egypt. Soc Sci Med 2003; 56:1837-1851.
2. Lutz W, O'Neill BC, Scherbov S. Demographics. Europe's population at a turning point. Science 2003; 299:1991-1992.
3. Grant J, Hoorens S, Sivadasan S, Loo MV, Davanzo J, Hale L, Butz W. Trends in European fertility: should Europe try to increase its fertility rate...or just manage the consequences? Int J Androl 2006; 29:17-24.
4. Skouby SO. Contraceptive use and behavior in the 21st century: a comprehensive study across five European countries. Eur J Contracept Reprod Health Care 2004; 9:57-68.
5. Benagiano G, Bastianelli C, Farris M. Contraception today. Ann N Y Acad Sci 2006; 1092:1-32.
6. Cibula D. Women's contraceptive practices and sexual behaviour in Europe. Eur J Contracept Reprod Health Care 2008; 13:362-375.
7. Skakkebaek NE, Jorgensen N, Main KM, Rajpert-De Meyts E, Leffers H, Andersson AM, Juul A, Carlsen E, Mortensen GK, Jensen TK, Toppari J. Is human fecundity declining? Int J Androl 2006; 29:2-11.
8. Jensen TK, Carlsen E, Jorgensen N, Berthelsen JG, Keiding N, Christensen K, Petersen JH, Knudsen LB, Skakkebaek NE. Poor semen quality may contribute to recent decline in fertility rates. Hum Reprod 2002; 17:1437-1440.
9. Nyboe Andersen A, Erb K. Register data on Assisted Reproductive Technology (ART) in Europe including a detailed description of ART in Denmark. Int J Androl 2006; 29:12-16.
10. Carlsen E, Giwercman A, Keiding N, Skakkebaek NE. Evidence for decreasing quality of semen during past 50 years. BMJ 1992; 305:609-613.
11. Auger J, Kunstmann JM, Czyglik F, Jouannet P. Decline in semen quality among fertile men in Paris during the past 20 years. N Engl J Med 1995; 332:281-285.
12. Swan SH, Elkin EP, Fenster L. Have sperm densities declined? A reanalysis of global trend data. Environ Health Perspect 1997; 105:1228-1232.
13. Swan SH, Elkin EP, Fenster L. The question of declining sperm density revisited: an analysis of 101 studies published 1934-1996. Environ Health Perspect 2000; 108:961-966.
14. Splingart C, Frapsauce C, Veau S, Barthelemy C, Royere D, Guerif F. Semen variation in a population of fertile donors: evaluation in a French centre over a 34-year period. Int J Androl 2011; 35:467-474.

15. Berling S, Wolner-Hanssen P. No evidence of deteriorating semen quality among men in infertile relationships during the last decade: a study of males from Southern Sweden. Hum Reprod 1997; 12:1002-1005.
16. Andolz P, Bielsa MA, Vila J. Evolution of semen quality in North-eastern Spain: a study in 22,759 infertile men over a 36 year period. Hum Reprod 1999; 14:731-735.
17. Irvine DS. Male reproductive health: cause for concern? Andrologia 2000; 32:195-208.
18. Skakkebaek NE, Rajpert-De Meyts E, Jorgensen N, Main KM, Leffers H, Andersson AM, Juul A, Jensen TK, Toppari J. Testicular cancer trends as 'whistle blowers' of testicular developmental problems in populations. Int J Androl 2007; 30:198-204.
19. Wong WY, Zielhuis GA, Thomas CM, Merkus HM, Steegers-Theunissen RP. New evidence of the influence of exogenous and endogenous factors on sperm count in man. Eur J Obstet Gynecol Reprod Biol 2003; 110:49-54.
20. Giwercman A, Giwercman YL. Environmental factors and testicular function. Best Pract Res Clin Endocrinol Metab 2011; 25:391-402.
21. Martini AC, Molina RI, Estofan D, Senestrari D, Fiol de Cuneo M, Ruiz RD. Effects of alcohol and cigarette consumption on human seminal quality. Fertil Steril 2004; 82:374-377.
22. Agarwal A, Desai NR, Ruffoli R, Carpi A. Lifestyle and testicular dysfunction: a brief update. Biomed Pharmacother 2008; 62:550-553.
23. Kolettis PN. Evaluation of the subfertile man. Am Fam Physician 2003; 67:2165-2172.
24. Bonde JP, Ernst E, Jensen TK, Hjollund NH, Kolstad H, Henriksen TB, Scheike T, Giwercman A, Olsen J, Skakkebaek NE. Relation between semen quality and fertility: a population-based study of 430 first-pregnancy planners. Lancet 1998; 352:1172-1177.
25. Zinaman MJ, Brown CC, Selevan SG, Clegg ED. Semen quality and human fertility: a prospective study with healthy couples. J Androl 2000; 21:145-153.
26. Guzick DS, Overstreet JW, Factor-Litvak P, Brazil CK, Nakajima ST, Coutifaris C, Carson SA, Cisneros P, Steinkampf MP, Hill JA, Xu D, Vogel DL. Sperm morphology, motility, and concentration in fertile and infertile men. N Engl J Med 2001; 345:1388-1393.
27. Slama R, Eustache F, Ducot B, Jensen TK, Jorgensen N, Horte A, Irvine S, Suominen J, Andersen AG, Auger J, Vierula M, Toppari J, Andersen AN, Keiding N, Skakkebaek NE, Spira A, Jouannet P. Time to pregnancy and semen parameters: a cross-sectional study among fertile couples from four European cities. Hum Reprod 2002; 17:503-515.
28. Barratt CL, Clements S, Kessopoulou E. Semen characteristics and fertility tests required for storage of spermatozoa. Hum Reprod 1998; 13 Suppl 2:1-7.
29. British Andrology S. British Andrology Society guidelines for the screening of semen donors for donor insemination (1999). Hum Reprod 1999; 14:1823-1826.
30. Ecochard R, Cottinet D, Mathieu C, Rabilloud M, Czyba JC. The mean of sperm parameters in semen donations from the same donor. An important prognostic factor in insemination. Int J Androl 1999; 22:163-172.
31. Carrell DT, Cartmill D, Jones KP, Hatasaka HH, Peterson CM. Prospective, randomized, blinded evaluation of donor semen quality provided by seven commercial sperm banks. Fertil Steril 2002; 78:16-21.
32. Kaufmann SJ, Eastaugh JL, Snowden S, Smye SW, Sharma V. The application of neural networks in predicting the outcome of in-vitro fertilization. Hum Reprod 1997; 12:1454-1457.
33. Kshirsagar AV, Murthy L, Chelu L, Lamb D, Ross L, Niederberger C. Predicting Outcomes for Intracytoplasmic Sperm Injection. Fertil Steril 2005; 84, Supplement 1:S274.

34. Parekattil SJ, Jr AJT. Artificial Intelligence Male Infertility Assistant. Fertil Steril 2005; 84, Supplement 1:S222.
35. Wald M, Sparks AE, Sandlow J, Van-Voorhis B, Syrop C, Niederberger CS. Computational Models for Prediction of IVF/ICSI Outcomes With Surgically Retrieved Sperm. Fertil Steril 2005; 84, Supplement 1:S12.
36. Ma Y, Chen B, Wang H, Hu K, Huang Y. Prediction of sperm retrieval in men with non-obstructive azoospermia using artificial neural networks: leptin is a good assistant diagnostic marker. Hum Reprod 2011; 26:294-298.
37. Lisboa PJ, Taktak AFG. The use of artificial neural networks in decision support in cancer: a systematic review. Neural Networks 2006; 19:408-415.
38. Haykin SS. Neural networks: a comprehensive foundation. London: Prentice-Hall; 1999: 842.
39. Kim CN, Yang KH, Kim J. Human decision-making behavior and modeling effects. Decis Support Syst 2008; 45:517-527.
40. Green M, Bjork J, Forberg J, Ekelund U, Edenbrandt L, Ohlsson M. Comparison between neural networks and multiple logistic regression to predict acute coronary syndrome in the emergency room. Artif Intell Med 2006; 38:305-318.
41. Kurt I, Ture M, Kurum AT. Comparing performances of logistic regression, classification and regression tree, and neural networks for predicting coronary artery disease. Expert Syst Appl 2008; 34:366-374.
42. Subashini TS, Ramalingam V, Palanivel S. Breast mass classification based on cytological patterns using RBFNN and SVM. Expert Syst Appl 2009; 36:5284-5290.
43. Polat K, Günes S. A new feature selection method on classification of medical datasets: Kernel F-score feature selection. Expert Syst Appl 2009; 36:10367-10373.
44. Gil D, Johnsson M, Garcia Chamizo JM, Paya AS, Fernandez DR. Application of artificial neural networks in the diagnosis of urological dysfunctions. Expert Syst Appl 2009; 36:5754-5760.
45. Gil D, Johnsson M. Using support vector machines in diagnoses of urological dysfunctions. Expert Syst Appl 2010; 37:4713-4718.
46. Gil D, Johnsson M, Garcia-Chamizo JM, Paya AS, Fernandez DR. Modelling of urological dysfunctions with neurological etiology by means of their centres involved. Applied Soft Computing 2011; 11:4448-4457.
47. Eustache F, Auger J, Cabrol D, Jouannet P. Are volunteers delivering semen samples in fertility studies a biased population? Hum Reprod 2004; 19:2831-2837.
48. Thorn P, Katzorke T, Daniels K. Semen donors in Germany: a study exploring motivations and attitudes. Hum Reprod 2008; 23:2415-2420.
49. World Health Organization. WHO laboratory manual for the examination and processing of human semen. Geneva: World Health Organization; 2010: 287.
50. Gil D, Girela JL, De Juan J, Gomez-Torres MJ, Johnsson M. Predicting seminal quality with artificial intelligence methods. Expert Syst Appl 2012; 39:12564-12573.
51. Ripley BD. Pattern recognition and neural networks. Cambridge: Cambridge university press; 1996: 403.
52. Bishop CM. Neural networks for pattern recognition. Oxford: Clarendon press; 1995: 482.
53. Rumelhart DE, Hinton GE, Williams RJ. Learning representations by back-propagating errors. Nature 1986; 323:533-536.

54. Pal M, University of Nottingham - GB. Factors Influencing the Accuracy of Remote Sensing Classification: A Comparative Study. University of Nottingham; 2002:
55. Hall M, Frank E, Holmes G, Pfahringer B, Reutemann P, Witten IH. The WEKA data mining software: an update. ACM SIGKDD Explorations Newsletter 2009; 11:10-18.
56. Kohavi R, Provost F. Glossary of terms. Mach Learning 1998; 30:271-274.
57. Sun Y, Kamel MS, Wong AKC, Wang Y. Cost-sensitive boosting for classification of imbalanced data. Pattern Recognit 2007; 40:3358-3378.
58. Conforti D, Guido R. Kernel based support vector machine via semidefinite programming: Application to medical diagnosis. Computers & Operations Research 2010; 37:1389-1394.
59. Kahramanli H, Allahverdi N. Design of a hybrid system for the diabetes and heart diseases. Expert Syst Appl 2008; 35:82.
60. Linneberg C, Salamon P, Svarer C, Hansen LK, Meyrowitsch J. Towards semen quality assessment using neural networks. In: Neural Networks for Signal Processing [1994] IV. Proceedings of the 1994 IEEE Workshop; 1994;
61. Samli MM, Dogan I. An artificial neural network for predicting the presence of spermatozoa in the testes of men with nonobstructive azoospermia. J Urol 2004; 171:2354-2357.
62. Li Y, Lin H, Li Y, Cao J. Association between socio-psycho-behavioral factors and male semen quality: systematic review and meta-analyses. Fertil Steril 2011; 95:116-123.
63. Almeida JS. Predictive non-linear modeling of complex data by artificial neural networks. Curr Opin Biotechnol 2002; 13:72-76.
64. Cleophas TJ, Cleophas TF. Artificial intelligence for diagnostic purposes: principles, procedures and limitations. Clin Chem Lab Med 2010; 48:159-165.
65. Jørgensen N, Andersen AG, Eustache F, Irvine DS, Suominen J, Petersen JH, Andersen AN, Auger J, Cawood EHH, Horte A. Regional differences in semen quality in Europe. Hum Reprod 2001; 16:1012-1019.
66. Carlsen E, Petersen JH, Andersson AM, Skakkebaek NE. Effects of ejaculatory frequency and season on variations in semen quality. Fertil Steril 2004; 82:358-366.
67. MacLEOD J. Effect of chickenpox and of pneumonia on semen quality. Fertil Steril 1951; 2:523-533.
68. Carlsen E, Andersson AM, Petersen JH, Skakkebæk NE. History of febrile illness and variation in semen quality. Hum Reprod 2003; 18:2089-2092.
69. Jung A, Schill WB, Schuppe HC. Genital heat stress in men of barren couples: a prospective evaluation by means of a questionnaire. Andrologia 2002; 34:349-355.
70. Künzle R, Mueller MD, Hänggi W, Birkhäuser MH, Drescher H, Bersinger NA. Semen quality of male smokers and nonsmokers in infertile couples. Fertil Steril 2003; 79:287-291.
71. Ramlau-Hansen C, Thulstrup AM, Aggerholm AS, Jensen MS, Toft G, Bonde JP. Is smoking a risk factor for decreased semen quality? A cross-sectional analysis. Hum Reprod 2007; 22:188-196.
72. Martini AC, Molina RI, Estofán D, Senestrari D, Fiol de Cuneo M, Ruiz RD. Effects of alcohol and cigarette consumption on human seminal quality. Fertil Steril 2004; 82:374-377.
73. Kort HI, Massey JB, Elsner CW, Mitchell-Leef D, Shapiro DB, Witt MA, Roudebush WE. Impact of body mass index values on sperm quantity and quality. J Androl 2006; 27:450-452.
74. Aggerholm AS, Thulstrup AM, Toft G, Ramlau-Hansen CH, Bonde JP. Is overweight a risk factor for reduced semen quality and altered serum sex hormone profile? Fertil Steril 2008; 90:619-626.

75. MacDonald A, Herbison G, Showell M, Farquhar C. The impact of body mass index on semen parameters and reproductive hormones in human males: a systematic review with meta-analysis. Hum Reprod Update 2010; 16:293-311.
76. Handelsman DJ, Conway AJ, Boylan LM, Turtle JR. Testicular function in potential sperm donors: normal ranges and the effects of smoking and varicocele. Int J Androl 1984; 7:369-382.
77. Rowley MJ, Leach DR, Warner GA, Heller CG. Effect of graded doses of ionizing radiation on the human testis. Radiat Res 1974; 59:665-678.
78. Oldereid NB, Rui H, Purvis K. Lifestyles of men in barren couples and their relationships to sperm quality. Eur J Obstet Gynecol Reprod Biol 1992; 43:51-57.
79. Mazurowski MA, Habas PA, Zurada JM, Lo JY, Baker JA, Tourassi GD. Training neural network classifiers for medical decision making: The effects of imbalanced datasets on classification performance. Neural Networks 2008; 21:427-436.

FIGURE LEGENDS

Figure 1. Diagram shows the proposed complementation of the Laboratory and the Artificial Intelligence approaches for the determination of the seminal profile.

Figure 2. The architecture of the MLP network (input layer, hidden layer and output layer). For a single neuron in the hidden layer it is shown the computation topology (Xi are inputs value; ωji is the connection weight between the input xi and the neuron j; θj is the threshold; fj is the activation function; yj is the output).

Figure 3. ROC curves showing the performance of the two neural networks developed (ANN1 and ANN2).

Table 1. Items in the original questionnaire about life habits and health status. Those marked are the ones chosen, after a preprocessing step, for the implementation of the artificial neural networks. (1) ANN1: To predict the sperm concentration. (2) ANN2: To predict the sperm motility.

Variables from original questionnaire

• Season in which the analysis was performed **(1)**	• Age at the time of analysis **(1,2)**	• Body mass index (BMI) **(2)**	• Marital status **(2)**
• Children	• Birth city (County)	• City of the family home (County)	• City of residence during the course (County)
• Number of siblings **(2)**	• Birth defect, genetic disorder or any hereditary disease	• childish diseases (ie , chicken pox, measles, mumps, polio) **(1)**	• Vaccines received **(2)**
• Allergy **(2)**	• Accident or serious trauma **(1)**	• Surgical intervention **(1)**	• Significant diseases
• High fevers in the last year **(1)**	• Chemotherapy	• Radiations (ie. x rays)	• Time since radiation exposure **(2)**
• Number of exposures	• Frequency of alcohol consumption **(1)**	• Alcohol consumption	• Smoking habit **(1)**
• Average number of cigarettes per day	• Years of exposure to cigarettes	• Coffee consumption	• Drugs consumption
• Number of hours spent sitting per day **(1)**	• Average hours sleep per day **(2)**	• Wearing tight clothes	• Warm baths **(2)**
• Average hours per week playing sports	• Average ejaculations per week		

Table 2. Characteristics and normalization ranges of the input parameters used in the two neural networks developed, one of them to predict the sperm concentration (ANN1) and the other one to predict the sperm motility (ANN2).

Seminal parameter predicted	Variable description	Values (min-max)	Normalized
Sperm concentration (ANN1)	Season in which the analysis was performed.	1) winter, 2) spring, 3) Summer, 4) fall.	(-1, -0.33, 0.33, 1)
	Age at the time of analysis.	18-36	(0 - 1)
	childish diseases (ie , chicken pox, measles, mumps, polio)	1) yes, 2) no.	(0, 1)
	Accident or serious trauma	1) yes, 2) no.	(0, 1)
	Surgical intervention	1) yes, 2) no.	(0, 1)
	High fevers in the last year	1) less than three months ago, 2) more than three months ago, 3) no.	(-1, 0, 1)
	Frequency of alcohol consumption	1) several times a day, 2) every day, 3) several times a week, 4) once a week, 5) hardly ever or never	(0, 1)
	Smoking habit	1) never, 2) occasional 3) daily.	(-1, 0, 1)
	Number of hours spent sitting per day	1-16	(0 - 1)
Sperm Motility (ANN2)	Age at the time of analysis.	18-36	(0 - 1)
	Body mass index (BMI)	17-34	(0 - 1)
	Marital status	1) unmarried and without a partner, 2) unmarried with partner, 3) married, 4) other.	(-1, -0.33, 0.33, 1)
	Number of siblings.	0-6	(0 - 1)
	Vaccines received	1) Children calendar, 2) Children calendar and tetanus, 3) only tetanus.	(-1, 0, 1)
	Allergy	1) yes, 2) no.	(0, 1)
	Time since radiation exposure	1) Less than 3 months, 2) More than 3 months.	(0, 1)
	Average hours sleep per day	6-12	(0 - 1)
	Warm baths	1) yes, 2) no.	(0, 1)

Table 3. Characteristics of the study population.

	Mean ± SD	5th – 95th percentile
Age	24,05 ± 4,38	18 - 33
Semen volume	3,85 ± 1,78	2 - 6,47
Sperm Concentration	67,68 ± 44,23	11 - 150
% of motile Sperm	54,17 ± 18,47	14,52 - 76,97

Table 4. Confusion matrix, showing the results of the classification process.

Sperm Concentration		Predicted Normal	Predicted Altered
actual	Normal	TP = 84	FN = 4
	Altered	FP = 6	TN = 6

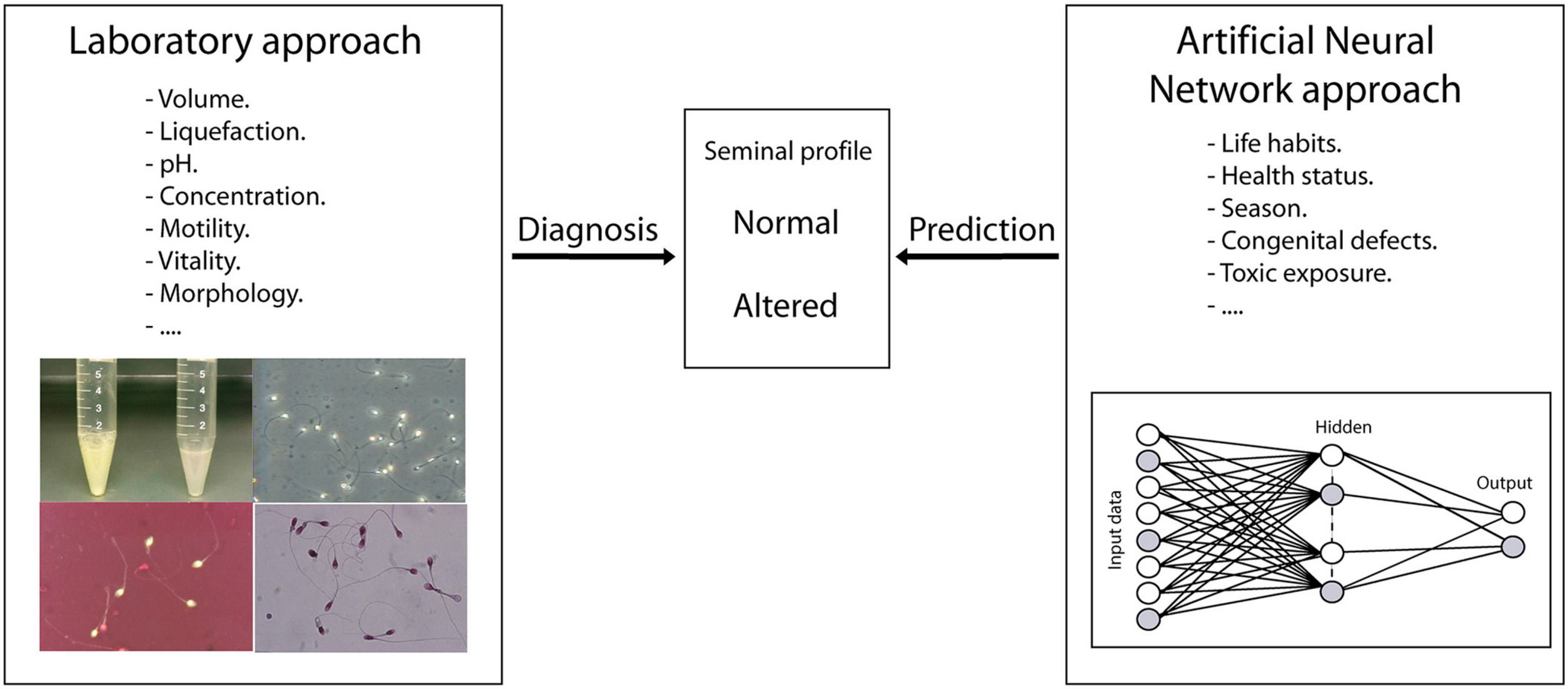
Sperm Motility		Predicted Normal	Predicted Altered
actual	Normal	TP = 75	FN = 9
	Altered	FP = 9	TN = 7

TP = True positive. Seminal parameter was normal and was correctly classified.
FP = False positive. Seminal parameter was altered and was incorrectly classified.
FN = False negative. Seminal parameter was normal and was incorrectly classified.
TN = True negative. Seminal parameter was altered and was correctly classified.

Table 5. Accuracy measures obtained by the MLP for each seminal parameter studied.

	Classification accuracy $\frac{TP + TN}{TP + FP + FN + TN} \times 100$	Sensitivity $\frac{TP}{TP + FN} \times 100$	Specificity $\frac{TN}{FP + TN} \times 100$	Positive predictive value $\frac{TP}{TP + FP} \times 100$	Negative predictive value $\frac{TN}{FN + TN} \times 100$
Sperm Concentration	90 %	95,45 %	50 %	93,33 %	60 %
Sperm Motility	82 %	89,29 %	43, 75 %	89,29%	43,75 %

Figure 1


Laboratory approach
- Volume.
- Liquefaction.
- pH.
- Concentration.
- Motility.
- Vitality.
- Morphology.
- ....
Diagnosis
Seminal profile
Normal
Altered
Prediction
Artificial Neural Network approach
- Life habits.
- Health status.
- Season.
- Congenital defects.
- Toxic exposure.
- ....
Input data
Hidden
Output

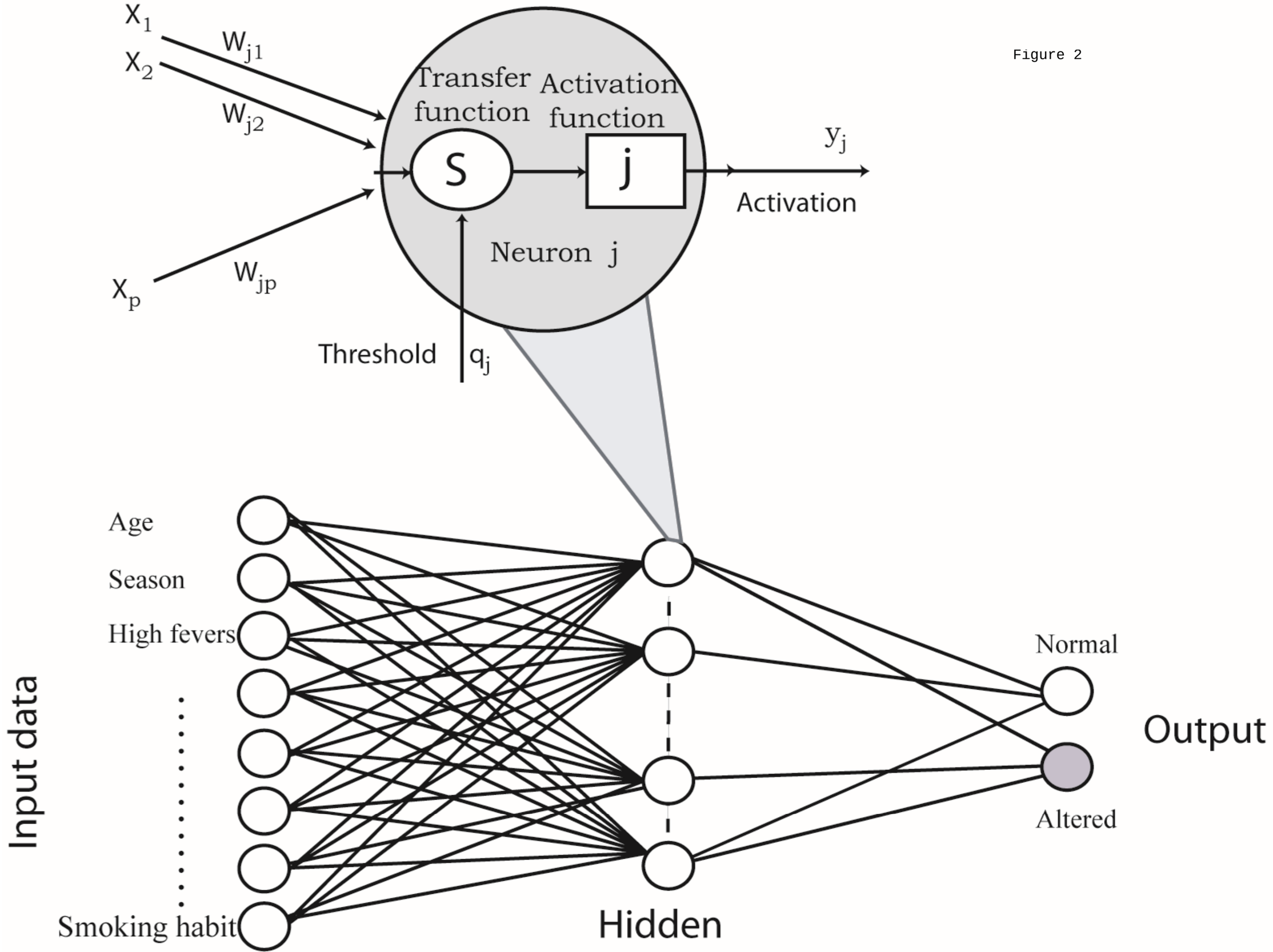

X1
X2
Xp
Wj1
Wj2
Wjp
Transfer function
Activation function
S
j
yj
Activation
Neuron j
Threshold
qj
Age
Season
High fevers
Smoking habit
Input data
Hidden
Normal
Altered
Output

Figure 2

Figure 3