Pergamon

0731-7085(95)01278-8

Journal of Pharmaceutical & Biomedical Analysis, Vol. 13, No. 3, pp. 205–211, 1995
Elsevier Science Ltd
Printed in Great Britain
0731-7085/95 $9.50 + 0.00

Classification of toxin-induced changes in ^{1}H NMR spectra of urine using an artificial neural network

MARIA L. ANTHONY,† VALERIE S. ROSE,‡ JEREMY K. NICHOLSON† and JOHN C. LINDON*‡

† *Department of Chemistry, Birkbeck College, University of London, Gordon House, 29 Gordon Square, London WC1H 0PP, UK*
‡ *Department of Physical Sciences, Wellcome Research Laboratories, Langley Court, Beckenham, Kent BR3 3BS, UK*

Abstract: NMR spectra of urine from rats treated with a range of liver, kidney and testicular toxins at various doses were measured and classified using neural network methods. Toxin-induced changes in the levels of 18 low molecular weight endogenous urinary metabolites were assessed using a simple semi-quantitative scoring system. These scores were used as input to an artificial neural network, the use of which has been explored as a means of predicting the class of toxin. With this limited data set, based only the level of the maximal changes of these 18 metabolites, the network was able to predict the class and hence target organ of the toxins. Renal cortical toxicity was well predicted as was liver toxicity. The few examples of renal medullary toxins in the data set resulted in relatively poor training of the network although correct classification was still possible.

Keywords: *^{1}H NMR; pattern recognition; urine; neural net; toxicity.*

Introduction

The usefulness of high field high resolution ^{1}H NMR spectroscopy of body fluids, principally urine, for the investigation of the biochemical changes that accompany toxin-induced organ damage is well documented [1]. The changes in the levels of endogenous metabolites found in urine following a toxic insult have been used to classify the site of the lesion and to explore the biological effects of the toxins [2–7]. The altered metabolite profile as detected using NMR spectroscopy can be regarded as providing a characteristic pattern for the type of toxin and this has led to the use of more formal computer-based methods for using the altered patterns to classify the urine samples according to the site of toxic lesion [7–9]. The methods used in general were dimension reduction techniques which allowed visualization of the toxin-related NMR spectra in simple two- or three-dimensional maps, thereby reducing the n-dimensional problem (where n is the number of metabolite intensities measured in each urine NMR spectrum) to one in which the similarity of the toxins can be visualized. This approach has been extended through the use of so-called "supervized" methods, whereby the sample class is used to determine which of the many measured descriptors are responsible for classifying the samples and which are irrelevant. This approach proved very successful for the evaluation of the time course of cellular degeneration and subsequent regeneration following an acute toxic insult [8]. For dimension reduction and visualization, the main techniques were cluster analysis, nonlinear mapping and principal components analysis [10].

This previous work demonstrated the usefulness and robustness of this type of pattern recognition method and showed the effectiveness of supervised methods for eliminating "metabolite noise", i.e. irrelevant descriptors from the data set. In the earlier work [7–9], this was achieved through the use of inter- and intra-group variance weight methods, although other techniques are available such as the method known as Soft Independent Modelling of Class Analogy, SIMCA [11], for classified data (such as for example, toxic vs non-toxic or liver vs kidney toxic) and partial least squares (PLS) [12] for continuous data. Such techniques, as well as leading to sample classifi-

* Author to whom correspondence should be addressed.

cation, can provide insight into the biochemical changes caused by the toxins. However, if this is not required and the need is for a robust automatic classification technique, then the application of artificial neural network software can be considered.

Neural network methodology [13] has found a rapidly increasing application in many areas of prediction both within and outside science. Within NMR spectroscopy, the method has already found use in the prediction of chemical shifts [14, 15] especially for carbohydrates [16–19] and for identifying cross-peaks in 2-D NMR [20]. A neural net purports to mimic the way in which the brain works. We have used a supervised network with back propagation and its mode of operation is represented schematically in Fig. 1. For each object, in this case an NMR spectrum derived from an animal dosed with a particular toxin, the n descriptors (here the levels of n metabolites) are entered at the input neurons for each sample and the network is trained by performing an optimization of weights of the interconnections between neurons, until, after input of all of the data of a training set, the values output at the output layer neurons are as close as possible to the values required. For example, if the desired output is a single number (in a drug metabolism study for example, this could be percentage recovery) then there will be one output neuron. In the case described here with five classes of sample there will be five output neurons and the correct values for a sample which falls into class 1 would be 1-0-0-0-0 and a correct value for a sample which falls into class 2 would be 0-1-0-0-0 and so on. The input to a neuron is the sum of the outputs of all of the neurons in the previous layer multiplied by the weight factor for each connection. It is these weights which are optimized at each stage of training the network. The output of a neuron is related to the summed inputs by a nonlinear transfer function, the commonest being of sigmoid shape. The optimization is therefore non-linear and consists of iteratively varying the weights until the output values for each sample are as close as possible to the target values. The theory of neural nets has been reviewed [13] and further details can be found therein.

We have now tested this neural network toxicological assessment approach by analysing the toxin induced changes in endogenous bio-

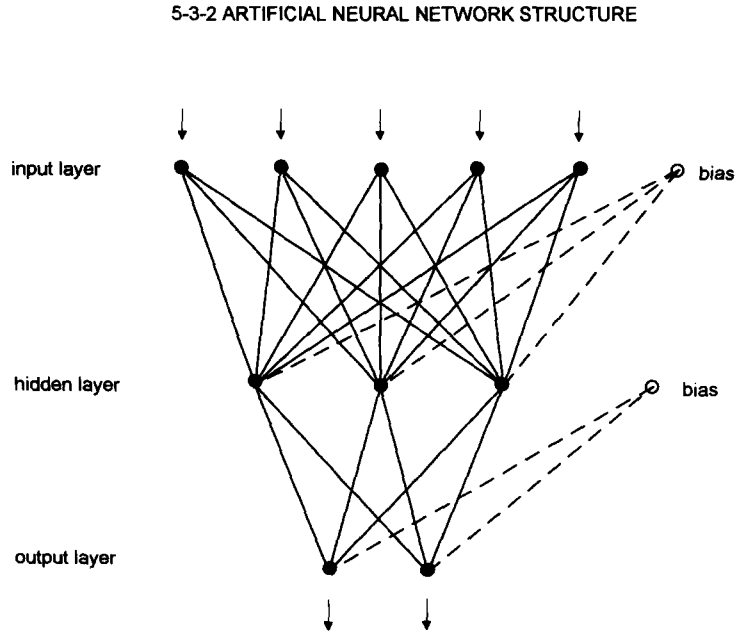

Figure 1
A representation of a typical artificial neural network shown here with a 5-3-2 architecture. It comprises an input layer of five neurons, one for each input descriptor, a hidden layer of three neurons, the number of which can be optimized according to the data, and an output layer of two neurons, one for each output variable. Bias neurons are included for the input and hidden layers. Input to a neuron, i, from a neuron in a previous layer, j, is given by $I_i = \Sigma w_{ji} \cdot O_j$, where O_j is the output from the previous neuron and the summation is over all nodes in the previous layer. w_{ji} is an adjustable weight for the i–j connection. Output from neuron i is governed by a sigmoidal transfer function of the form $O_i = (1 + \exp(-I_i))^{-1}$.

chemicals in urine as measured using ^{1}H NMR spectroscopy to ascertain whether the methods provide a robust approach which could lead to automatic toxin classification.

Experimental

The details of the toxicological experiments, NMR spectroscopic measurements and data collection have been described earlier [2–7]. All toxins used produced characteristic target organ-specific lesions except the lowest doses of cadmium (6 μmol kg^{-1}) and mercury (0.5 mg kg^{-1}) which were below the toxic threshold [2, 3]. The changes in the levels of 18 endogenous metabolites were assessed for 25 toxins over a 48-h period after dosing using a simple scoring system as described previously [7] and the results are given in Table 1. These relate to the maximum change irrespective of which time point in the toxicological study the effect occurred. Each score is the modal value for a group of animals, where positive values indicate an increase over control, zero represents an unchanged level and negative values denote decreased values from control. The complete sample set consisted of 25 samples with 12, 2, 5, 2 and 4 members of classes 1–5 consisted of 25 samples with 12, 2, 5, 2 and 4 members of classes 1–5, respectively. The small number of samples available in most classes precluded the use of rigorous testing of significance levels and predictability by cross-validation. Cross-validation is inappropriate for small sample sizes as there is little redundancy in the data and each point contributes significantly to the model. In this situation, one can best test a model by predicting the class of new samples. To this end, a small test set was extracted from the sample set using three samples from the two largest classes. This comprised two class-1 toxins and one class-3 toxin which were combined with three further class-1 toxin data sets which became available to form a six sample test set. No members of classes 2, 4 and 5 were represented in the test set due to paucity of data.

The sample classes were as described in Table 1, i.e. five in all, including a non-toxic class from only two samples (resulting from non-toxic dose levels). The classification of the samples was confirmed by histopathology [2–7] and the classes were renal cortical (class-1), renal medullary (class-2), hepatic (class-3), non-toxic (class-4) and testicular/reproductive (class-5). The toxins were classified according to target organ based on the histopathology results [2–7] and each was given five values, thus a cortical toxin would have values 1,0,0,0,0 and a renal medullary toxins had values 0,1,0,0,0, a hepatic toxin was 0,0,1,0,0, a control substance or a non-toxic level had 0,0,0,1,0 and reproductive organ toxins were categorized as 0,0,0,0,1. Hence, a perfectly trained network should return such values for perfect classification. Any deviation from 1 or 0 reflects error in the classification process.

Neural net calculations were performed on a 386 Compaq PC fitted with a Delta II floating point processor from SAIC [21] using the SAIC ANSim software package [21]. The 18 biochemical descriptors given in Table 1 were range-scaled in a linear fashion so that all values were in the range −0.5 to +0.5 prior to input to the network. The inter-neuron weights were initially set to random values in the range −0.3 to +0.3 and the network was trained using a learning rate of 0.1 and a momentum term of 0.6 for both layers. The network geometry comprised 18 input neurons, 9 hidden neurons and 5 output neurons with bias neurons in the first two layers. The bias neurons are introduced to off-set the origin of the transfer function. A bias neuron has an output value of one and an associated weight. The bias neuron-to-neuron weights are trained in a similar way to the other inter-neuron weights.

Results

Training of the network

The network was trained for 100 cycles by the input of 18 descriptors for 19 toxins using the back propagation method. This represented the minimum amount of training required to give good classification for the majority of the training set. Further training failed to significantly improve classification and could result in over-fitting problems and so reduce predictability of the network outside the training set.

Prediction of individual toxin classes based on overall trained network

Table 2 shows the predicted class values for the toxins used in the training and test sets. The network models all of the renal cortical toxins correctly with output values >0.89 in all cases, both for those in the training set and for

Table 1
Metabolite level changes from the ^{1}H NMR spectra of urine following a single dose of various toxins

Toxin	ace	ala	HB	cn	cit	glc	gln	20G	hip
Hg (0.5)*	−1	0	2	0	−2	0	0	−2	−1
Hg (1.0)*	1	3	3	0	−2	2	2	2	0
Hg (2.0)*	3	3	3	−2	−3	3	2	3	1
Hg (2.0 fed)*	1	1	0	−1	−3	2	1	−2	−1
PAP (50)*	1	2	1	−1	−1	2	1	−1	−1
PAP (100)*	1	2	1	−1	−2	3	2	−1	−1
HCBD*	1	2	1	0	0	3	2	−1	−3
UN*	1	0	3	−2	−3	3	2	−2	−3
PI*	3	1	0	0	1	1	0	−2	0
BEA*	2	1	0	0	−2	1	0	−3	−1
HYD*	3	0	0	0	−2	0	0	−2	0
CC14*	0	0	0	−1	−2	0	0	−2	−1
THIO*	0	0	0	−1	−2	0	0	0	−2
ALY*	3	1	0	0	−1	0	0	−2	0
Cd6*	0	0	0	0	0	0	0	0	−1
Cd9*	0	0	0	0	−2	0	0	−1	−1
Cd12*	0	0	1	0	−2	0	0	−1	−1
Cd24*	1	0	1	0	−2	0	0	−1	−2
Cd(F)*	0	0	0	0	−2	0	0	−2	−1
CrO4†	1	1	0	−1	0	3	2	−1	−2
Hg (1.5)†	2	3	3	−2	−3	3	2	2	−1
CPH†	2	2	3	−2	−2	3	2	−2	−3
TCTFP†	3	2	3	−2	−2	3	2	−2	−2
DCVHC†	3	2	2	−2	−2	3	2	−3	−3
ANIT†	2	0	0	0	−2	1	0	−2	0

Abbreviations: ace — acetate, ala — alanine, HB — β-hydroxybutyrate, cn — creatinine, cit — citrate, glc — glucose, gln — glutamine, 2-OG — 2-oxoglutarate, hip — hippurate, lac — lactate, suc — succinate, TMAO — trimethylamine-N-oxide, val — valine, DMA — dimethylamine, lys — lysine, tau — taurine, DMG — dimethylglycine, cr — creatine, CrO4 — sodium chromate, ANIT — α-naphthylisothiocyanate, CPH — cephaloridine; TCTFP — 3,3,3-trifluorotrichloroprop-1-ene, DCVHC — dichlorovinyl-homocystine.

Toxins: Hg — mercury (II) chloride, PAP — para-aminophenol, HCBD — hexachlorobutadiene, PI — propyleneimine, BEA — bromoethanamine hydrobromide, HYD — hydrazine, CC14 — carbon tetrachloride, THIO — thioacetamide, ALY — allyl alcohol, UN — uranyl nitrate, Cd — cadmium, F — female.

Table 2
Output values from neural network analysis

Toxin	Cortical	Medullary	Hepatic	Non-toxic	Reproductive
Hg (fed)*	0.895	0.105	0.087	0.079	0.012
Hg (0.5)*	0.104	0.123	0.128	0.199	0.245
Hg (1.0)*	0.989	0.084	0.032	0.079	0.012
Hg (2.0 fed)*	0.980	0.060	0.048	0.064	0.005
PAP (50)*	0.893	0.122	0.072	0.081	0.012
PAP (100)*	0.973	0.094	0.038	0.064	0.006
UN*	0.971	0.061	0.037	0.070	0.009
HCBD*	0.968	0.110	0.030	0.065	0.007
PI*	0.146	0.373	0.191	0.121	0.081
BEA*	0.071	0.371	0.078	0.124	0.331
HYD*	0.051	0.151	0.874	0.156	0.061
CC14*	0.053	0.130	0.847	0.170	0.067
THIO*	0.055	0.125	0.834	0.172	0.070
ALY*	0.097	0.190	0.720	0.138	0.049
Cd6*	0.071	0.187	0.210	0.184	0.208
Cd9*	0.012	0.196	0.060	0.172	0.827
Cd12*	0.021	0.179	0.050	0.175	0.823
Cd24*	0.016	0.182	0.101	0.170	0.797
Cd(F)*	0.028	0.187	0.082	0.173	0.686
CrO4†	0.919	0.120	0.046	0.075	0.013
Hg (1.5)†	0.979	0.053	0.046	0.066	0.006
CPH†	0.979	0.059	0.075	0.066	0.004
TCTFP†	0.983	0.079	0.029	0.061	0.005
DCVHC†	0.983	0.079	0.026	0.060	0.006
ANIT†	0.058	0.172	0.849	0.150	0.054

*Training set.
†Test set.
A perfect classification would result in a single value of 1.0 and four values of 0.0 for each toxin.

lac	suc	TMAO	val	DMA	lys	tau	DMG	cr	class
−2	−1	0	0	0	0	0	0	0	4
2	1	2	2	0	0	0	0	0	1
3	3	3	3	−2	1	0	0	0	1
3	−3	−2	1	−1	1	0	1	0	1
2	1	0	1	0	1	0	0	0	1
3	1	0	2	0	2	0	1	0	1
3	1	0	2	0	2	0	−1	0	1
2	−2	−1	1	−2	1	0	0	0	1
1	3	−3	0	2	0	0	3	0	2
1	2	−3	0	2	0	0	3	2	2
2	−2	0	0	−1	0	3	0	0	3
0	−2	−1	0	−1	0	3	0	0	3
0	−2	−1	0	−1	0	3	0	0	3
2	3	−2	0	0	0	1	−3	0	3
0	−1	0	0	0	0	0	0	0	4
0	−1	0	0	0	0	0	0	3	5
0	−1	0	0	0	0	0	0	3	5
0	−2	0	0	0	0	1	0	3	5
0	−2	0	0	0	0	0	0	2	5
1	−1	0	0	0	1	0	0	0	1
2	1	2	2	−2	1	0	−2	0	1
3	−1	0	2	−3	2	0	1	0	1
3	2	−2	2	−2	2	0	−1	0	1
3	−2	−2	2	−2	2	0	−1	0	1
1	2	0	0	0	0	3	0	0	3

Doses: Hg — 0.5–2.0 mg kg^{-1}; PAP — 50 and 100 mg kg^{-1}; HCBD — 200 mg kg^{-1}; PI — 20 μl kg^{-1}; BEA — 250 mg kg^{-1}; HYD — 60 mg kg^{-1}; CC14 — 2 ml kg^{-1}; THIO — 20 mg kg^{-1}; ALY — 50 μl kg^{-1}; UN — 10 mg kg^{-1}; Cd — 6–24 μmol kg^{-1}; sodium chromate — 20 mg kg^{-1}; ANIT — 300 mg kg^{-1}; CPH — 750 mg kg^{-1}; TCTFP — 40 mg kg^{-1}; DCVHC — 40 mg kg^{-1}.

*Training set.

†Test set.

Table 3
Predicted toxin classes

Toxin	Observed class	Predicted class*	Predicted class†
Hg (0.5)	Non-toxic‡	Reproductive	Incorrect
Hg (1.0)	Cortical	Cortical	Correct
Hg (1.5)	Cortical	Cortical	Correct
Hg (2.0)	Cortical	Cortical	Correct
Hg (2.0 fed)	Cortical	Cortical	Correct
PAP (50)	Cortical	Cortical	Correct
PAP (100)	Cortical	Cortical	Correct
CrO4	Cortical	Cortical	Correct
UN	Cortical	Cortical	Correct
HCBD	Cortical	Cortical	Correct
CPH	Cortical	Cortical	Correct
TCTFP	Cortical	Cortical	Correct
DCVHC	Cortical	Cortical	Correct
PI	Medullary	Medullary	Unclassified
BEA	Medullary	Medullary	Unclassified
HYD	Hepatic	Hepatic	Correct
CCl4	Hepatic	Hepatic	Correct
THIO	Hepatic	Hepatic	Correct
ALY	Hepatic	Hepatic	Correct
ANIT	Hepatic	Hepatic	Correct
Cd6	Non-toxic‡	Hepatic/reproductive	Incorrect
Cd9	Reproductive	Reproductive	Correct
Cd12	Reproductive	Reproductive	Correct
Cd24	Reproductive	Reproductive	Correct
Cd (F)	Reproductive	Reproductive	Correct

*Predicted on the basis of the highest output score ("winner-takes-all").

†Predicted on the basis of an output score >0.67 correct, 0.33–0.67 unknown, <0.33 incorrect.

‡Below the toxic threshold for the compound, producing no pathological abnormalities or changes in urinary enzyme excretion.

those applied to the trained network in the test set. Hg (0.5) represents a non-toxic dose and the network correctly does not classify this as a cortical toxin. However, it also does not give it a high value as a control sample. The network also classifies all of the liver toxins correctly, all output values being high, the lowest is ALY at 0.72. Again in the test set, the toxin ANIT is correctly predicted to be a hepatotoxin. In addition, all of the reproductive system toxins in the training set are classified correctly with output values >0.80 except for Cd(F), the dose of cadmium to female rats, which has a lower output value of 0.69, probably as a result of the smaller observed perturbation of urinary metabolite levels in female animals. The non-toxic low dose of cadmium Cd6 is correctly not classified as toxic although again the output value for a control class is also low. The ability to classify the renal medullary toxins is not so good although the output values for the correct class are much higher than for the incorrect classes. The small number of samples is a difficulty in the training of the network. On a "winner-takes-all" basis the samples in both the training and test sets are largely correctly classified (Table 3). If this approach is applied, then correct classification is achieved for all samples except for the non-toxic Hg (0.5) sample which is classified as a testicular toxin and the non-toxic Cd6 dose which is classified equally as a renal cortical or testicular toxin. The results are summarized in Table 3.

Taking a more stringent criterion of a value of >0.67 being correct, values between 0.33 and 0.67 taken as unclassified and values <0.33 as being incorrect also results in good classification as shown in Table 3. Here only the two non-toxic dose levels are classified incorrectly and the two renal medullary toxins are classified as unknown.

Conclusions

The neural network approach to sample classification is in general predictive of the sample class. It appears to be reasonably robust and once the network is trained the prediction of new samples is rapid and automatic. It is hoped to more finely tune the network with respect to improving its predictive power and optimizing its architecture (i.e. the number of nodes in the hidden layer) when more data becomes available and rigorous significance testing becomes feasible. However, the principal disadvantage is common to all neural network studies in that it is difficult to ascertain from the network which of the original sample descriptors are responsible for the classification. We have also used other pattern recognition methods to classify the urine NMR spectra from a more limited set of nephrotoxins alone [9]. The approach should be of general and widespread applicability and offers promise in areas of NMR of biofluids as diverse as toxicology of experimental drugs and in clinical diagnosis where pattern recognition methods based upon NMR spectroscopy of urine have been used to classify samples from patients with inborn errors of metabolism [22].

References

[1] J.K. Nicholson and I.D. Wilson, *Prog. NMR Spectrosc.* **21**, 449–501 (1989).
[2] J.K. Nicholson, J.A. Timbrell and P.J. Sadler, *Mol. Pharmacol.* **27**, 644–651 (1985).
[3] J.K. Nicholson, D.P. Higham, J.A. Timbrell and P.J. Sadler, *Mol. Pharmacol.* **36**, 398–404 (1989).
[4] K.P.R. Gartland, F.W. Bonner, J.A. Timbrell and J.K. Nicholson, *Arch. Toxicol.* **63**, 97–106 (1989).
[5] K.P.R. Gartland, F.W. Bonner and J.K. Nicholson, *Mol. Pharmacol.* **35**, 242–250 (1988).
[6] M.L. Anthony, K.P.R. Gartland, C.R. Beddell, J.C. Lindon and J.K. Nicholson, *Arch. Toxicol.* **68**, 43–53 (1994).
[7] K.P.R. Gartland, C.R. Beddell, J.C. Lindon and J.K. Nicholson, *Mol. Pharmacol.* **39**, 629–642 (1991).
[8] E. Holmes, F.W. Bonner, B.C. Sweatman, J.C. Lindon, C.R. Beddell, E. Rahr and J.K. Nicholson, *Mol. Pharmacol.* **42**, 922–930 (1992).
[9] M.L. Anthony, B.C. Sweatman, C.R. Beddell, J.C. Lindon and J.K. Nicholson, *Mol. Pharmacol.* **46**, 199–211 (1994).
[10] M.A. Sharaf, D.L. Illman and B.R. Kowalski, *Chemometrics*. J. Wiley & Son, New York (1986).
[11] S. Wold, *Pattern Recognition* **8**, 127–139 (1976).
[12] P. Geladi and B.R. Kowalski, *Anal. Chim. Acta* **185**, 1–17 (1986).
[13] P.D. Wasserman, *Neural Computing: Theory and Practice*. Van Nostrand Reinhold, New York (1989).
[14] L.S. Anker and P.C. Jurs, *Anal. Chem.* **64**, 1157–1164 (1992).
[15] G.M.J. West, *J. Chem. Inf. Comput. Sci.* **33**, 577–589 (1993).
[16] B. Meyer, T. Hansen, D. Nute, P. Albersheim, A. Darvill, W. York and J. Sellers, *Science* **251**, 542–544 (1991).
[17] J.W. Ball and P.C. Jurs, *Anal. Chem.* **65**, 3615–3621 (1993).
[18] W.S. Goux and D.S. Weber, *Carbohydr. Res.* **240**, 57–69 (1993).
[19] J.U. Thomsen and B. Meyer, *J. Mag. Reson.* **84**, 212–217 (1989).
[20] B.J. Hare and J.H. Prestegard, *J. Biomol. NMR* **4**, 35–46 (1994).

[21] SAIC ANSim, SAIC, 102260 Campus Point Drive, MS71, San Diego, CA, USA.
[22] E. Holmes, P.J.D. Foxall, J.K. Nicholson, G.H. Neild, S.M. Brown, C.R. Beddell, B.C. Sweatman, E. Rahr, J.C. Lindon, M. Spraul and P. Neidig, *Anal. Biochem.* **220**, 284–296 (1994).

[*Received for review 27 September 1994; revised manuscript received 7 November 1994*]