Feasibility of Developing a Neural Network for Prediction of Human Pharmacokinetic Parameters from Animal Data

Ajaz S. Hussain,[1,2] Robert D. Johnson,[1] Nimish N. Vachharajani,[1] and Wolfgang A. Ritschel[1]

Received April 3, 1992; accepted September 22, 1992
KEY WORDS: neural computing; prediction; pharmacokinetics; humans; animals; allometry.

INTRODUCTION

The prediction of pharmacokinetic parameters in man from data obtained in laboratory animals (interspecies scaling) is most commonly achieved through the allometric approach. Allometry attempts to establish a quantitative relationship between the pharmacokinetic parameters of interest and a physiological characteristic of the animal such as body weight, brain weight, liver weight, or body surface area (1–6). Although empirical, allometric scaling relies on the observation that many mammalian physiological processes are highly correlated with body weight. Therefore, it is reasonable to infer that drug distribution and elimination may also be related to animal size. The basic equation used in allometric species scaling is given by

$$PP = \alpha \cdot X^{\beta} \quad (1)$$

where PP is the pharmacokinetic parameter to be scaled, X is the animal physiologic parameter, α is the allometric coefficient, and β is the allometric exponent.

The allometric equation, which assumes that the pharmacokinetic parameters in mammals can be mapped to a single physiologic parameter by a log-linear transformation, is an example of a one-dimensional transformation. In several instances, especially for drugs metabolized by the mixed function oxidase system or low-extraction ratio drugs, the human parameters do not usually coincide with the animal data following this one-dimensional transformation. However, if the pharmacokinetic parameter (typically clearance) is corrected for the maximum life span potential (MLP), the human data are brought into line with the other species (3). The MLP, however, is a function of both body weight and brain weight; therefore, the simple one-dimensional transformation is no longer an adequate mapping function. Nevertheless, several retrospective studies have demonstrated the utility of the allometric approach (7–15).

In the allometric scaling approach, the calculations are performed essentially in isolation, i.e., only the animal data for a particular drug are used in the calculation, without using the available information from other drugs and/or structural information of the molecule. There may, however, exist some relevant information in other data sets that, if utilized, may improve our ability to predict human data. Data evaluation with a suitable nonlinear pattern recognition system may also offer some advantages. Artificial neural networks (ANNs) have been reported to be useful pattern recognition tools for a wide variety of problems in chemistry and pharmacy (16 (review),17,18). Some of these studies have compared ANN with common statistical techniques such as multiple linear or polynomial regression analysis, nearest-neighbor classifier, maximum-likelihood estimation, and Bayesian estimation. The performance of ANNs was reported to be comparable or superior to that of these techniques. The advantage of ANNs over statistical estimation techniques is that no a priori knowledge of the underlying statistical nature of the problem is required and no simplifying assumptions need to be made for application of this technique in a sparse data environment (17). In this communication, we have examined the feasibility of developing an ANN tool for the prediction of the apparent volume of distribution (V_z) and the total clearance (Cl_{tot}) in humans from a set of animal data alone and from a combination of animal data and structural descriptors based on chemical graph theory.

METHODS

Pharmacokinetic Data and Allometric Calculations

The PPs (Cl_{tot} and V_z) of 14 drugs were all taken from the literature (7–15,19–26). Only commonly used laboratory animals (total of four different species) such as the mouse, rat, rabbit, monkey, and dog were used. For most of these drugs, the allometric approach has been shown to be successful. For two drugs (DDC and AZT), the human PPs were obtained in patients, not in healthy volunteers. Also, for these drugs, the steady-state volume of distribution (V_{ss}) was used. Inclusion of these drugs assumes that interspecies variability is much greater than intraspecies variability. The allometric parameters [α and β from Eq. (1)] for the PPs, using the body weights as the physiologic variable, were estimated by linear regression analysis after log-log transformation. The resultant allometric equation [Eq. (1)] was then used to predict the results in man. In addition, for Cl_{tot}, the MLP correction (Cl_{tot} * MLP) was also used.

Chemical Graph Theory-Based Molecular Descriptors

The valence chi (χ) and the kappa shape indexes (κ_α) proposed by Hall and Kier (27) were calculated by the Molconn-χ program (Hall Associates Consulting, Quincy, MA).

ANN Analysis

The ANN analysis was performed with the NeuralWorks Professional II/Plus software (NeuralWare, Inc., Pittsburgh, PA) on an IBM-compatible 486DX/33mHz computer. A three-layer (input, hidden, and output) nonlinear feedforward network with the hyperbolic tangent transfer function based on the Extended-Delta-Bar-Delta (EDBD) al-

[1] Division of Pharmaceutics and Drug Delivery Systems, College of Pharmacy, University of Cincinnati Medical Center, Cincinnati, Ohio 45267-004.
[2] To whom correspondence should be addressed.

0724-8741/93/0300-0466$07.00/0

gorithm (28) was used. This algorithm is a modified version of the standard backpropagation algorithm and reduces the training time by using individual (for each node) time varying momentum and learning rate coefficients that are adjusted heuristically during training.

Input/Output Data Format

The following input/output formats were investigated for ANN analysis based on the animal data.

1. BW_1,BR_1,MLP_1,PP_1,.....,BW_4,BR_4,MLP_4,BW_m, BR_m,MLP_m/PP_m
2. $(PP/BW)_1$,.....,$(PP/BW)_4/(PP/BW)_m$
3. $(Cltot * MLP/BW)_1$,.., $(Cltot * MLP/BW)_4/(Cltot * MLP/BW)_m$
4. $(Cltot/BW)_1$, $(Cltot * MLP/BW)_1$,.., $(Cltot/BW)_4$, $(Cltot * MLP/BW)_4/(Cltot * MLP/BW)_m$

where BW is body weight, BR is brain weight, and the subscript numbers refer to the individual species, and m refers to man. In addition, the logarithmic transformation of the above parameters were also evaluated. The arrangement of the data set was in the increasing order of the body weight (see Table I).

Min–Max Values. The input/output data were scaled to the -0.5 to $+0.5$ range for V_z and to the -1 to $+1$ range for Cl_{tot} by the following equation:

$$\text{scaled value} = \gamma_1 + (x - x_1) \cdot \left(\frac{x_2 - x_1}{\gamma_2 - \gamma_1}\right) \qquad (2)$$

where x is the value to be scaled, x_1 is the unscaled low value (0), x_2 is the unscaled high value, γ_1 is the scaled low value (-1 or -0.5), and γ_2 is the scaled high value (1.0 or 0.5).

ANN Training. The number of hidden layer nodes was selected by training the network with 1, 2, 3, and 4 hidden nodes and selecting the lowest number that resulted in a smooth and stable reduction in the root mean square error (RMS) as a function of training sequence. The network was trained for 1000 epochs (number of training examples). The connection weights were initially randomized to values in the -0.1 and $+0.1$ range. The ANN's performance was evaluated by the "one-out" method, i.e., one data set (one drug) was excluded during training, after which the trained ANN was evaluated for its ability to predict the PPs of the excluded drug. This was repeated until all drugs were excluded once from the training set. The ANN-predicted Cl_{tot} and V_z were then compared with the predictions obtained from the standard allometric techniques.

For the ANN containing both animal data and structural descriptors, input/output format 2 was used for V_z and format 3 was used for Cl_{tot}. The valence path indices of order 1 through 3 ($^1\chi_v$, $^2\chi_v$, $^3\chi_v$), the third-order cluster ($^3\chi_{cl}$), the fourth-order path/cluster ($^4\chi_{p/cl}$), and shape indices of order 1 through 3 ($^1\kappa_\alpha$, $^2\kappa_\alpha$, $^3\kappa_\alpha$) were used. Of these, the most relevant structural descriptors were selected by training the network on the entire set of inputs (all structural descriptors and animal data). After training, each input node was sequentially and individually disabled (set to zero) to test its contribution (increase in RMS). Those input nodes that did not give a significant increase in the RMS were eliminated. The resultant network was used for the PP prediction as described above.

RESULTS AND DISCUSSION

Several types of neural networks, each differing in their training algorithm and architecture, have been proposed. The delta backpropagation network (16) is the most widely used, and its training algorithm is, in some aspects, similar to (multiple) nonlinear regression analysis. In this investigation, we used the EDBD training algorithm (28) because, in previous studies, we have found it to be more efficient (shorter training time) compared to the simple delta backpropagation algorithm (29).

Several different input/output (animal) data formats were investigated. Our initial efforts were focused on format 1, in which individual body weights, brain weights, MLPs, and PPs were used. Both training and generalization were successful; however, use of this format was discontinued because data for all drugs in the same species were not available. Use of parameters of different species for the same

Table I. Results for Volume of Distribution (V_z)[a]

Drug	Species	Observed	Allometry	ANN I	ANN II
DDC (10)	A,B,D,E	0.54	1.61	0.53	1.03
SM-1652 (19)	A,B,C,F	0.12	0.16	0.65	0.23
Phencyclidine (11)	A,B,E,F	6.19	12.9	6.14	6.49
FCE-22101 (12)	B,C,E,F	0.44	0.51	0.57	0.72
Cefotetan (20,21)	A,B,C,F	0.18	0.20	0.55	0.25
Erythromycin (9)	A,B,C,F	1.04	4.11	2.14	0.88
Methotrexate (7)	A,B,E,F	0.65	0.66	0.59	0.65
Cyclophosphamide (22)	A,B,E,F	1.14	0.93	0.63	1.02
Ceftizoxime (23,24)	A,B,E,F	0.39	0.10	0.49	0.14
Caffeine (8)	A,B,C,E	0.73	0.92	0.64	0.85
Propafenone (13)	A,B,C,F	1.72	3.48	1.27	1.65
AZT (14)	A,B,E,F	1.40	1.24	0.68	0.72
CAS-94457-09-7 (25,26)	A,B,C,F	0.72	0.42	0.93	0.84
Coumarin (15)	B,C,E,F	2.21	11.6	4.43	1.94
RMS			0.87	0.20	0.08

[a] A = mouse; B = rat; C = rabbit; D = cat; E = monkey; F = dog. V_z expressed as L/kg. ANN I corresponds to the predicted V_z without topological indices. ANN II corresponds to the predicted V_z with topological indices. RMS is the root mean square error.

input node may yield a pattern (in the data set) quite different from the one we hope to extract. For example, data points that have different species at a particular input node (i.e., rat instead of mouse) may be grouped differently by the network. Normalizing the PP to the physiologic parameters (BW and/or MLP) reduces the number of inputs and reduces the chance of undesirable patterns in the data set. Ideally, data from the same animal species should be used.

For the prediction of V_z, input/output format 2 using both numeric and log-transformed data was evaluated. The numeric format (test set RMS = 0.2) appeared to be superior to the log-format (RMS = 0.48); however, a min–max range of greater than −0.5 to 0.5 (−1.0 to 1.0 is the typical range for tanh transfer function) resulted in network paralysis during a few (one-out) training sets. In this communication, only the numeric format results are listed. Two hidden layer nodes were sufficient to achieve a a RMS of 0.09 and 0.04 for ANN I and ANN II, respectively. Results of the allometric and the ANN (I) predictions (one-out test method) are shown in Table I. The overall prediction error for the ANN approach was lower (RMS = 0.2) than that of the allometric approach (RMS 0.87). However, ANN I was not sensitive for low V_z values (less than 0.5). Additional hidden nodes (up to 4) did not improve the prediction of low V_z's. Use of the smallest possible volume of distribution for a hypothetical drug (plasma volume) as an additional input value in the training set did improve the predictions slightly (data not shown).

Additional input parameters such as extent of protein binding, partition coefficient, and ionization constant would certainly be relevant for this problem; however, these data were not available for all the drugs used in this study. Chemical graph theory has been developed to characterize the topology of a molecule in terms of numerical indices. These indices have been shown to be related to several physicochemical properties (solubility, partition coefficient, etc.) and the pharmacologic/toxicologic activity of several compounds (27). Application of chemical graph theory for quantitative structure–pharmacokinetic relationships of a series of benzodiazepines has also been reported (30). Therefore, we attempted to supplement the animal data with topological indices using the χ_v and κ_α indices to develop an enhanced network (ANN II). From the relevant feature extraction procedure, the $^1\chi_v$, $^4\chi_{p/cl}$, $^1\kappa_\alpha$, $^2\kappa_\alpha$, and $^3\kappa_\alpha$ were found to complement the animal data as noted by the reduced RMS (0.08). These results suggest that knowledge of only V_z in animals may not be sufficient to predict the value in humans for all drugs. Since V_z is a function of the extent of binding to blood constituents and, in many cases, the extent of binding differs among the various species, the improved RMS may be a result of a relationship between the structural descriptors and the binding characteristics. However, this hypothesis remains to be verified.

Results for the Cl_{tot} are summarized in Table II. The predicted values listed under Allometry MLP Corr. were obtained by indiscriminate use of this correction factor. Application of the MLP correction for SM-1652, phencyclidine, erythromycin, cyclophosphamide, and propafenone improved the prediction over that of simple allometry. For caffeine and CAS-94457-09-7, both of which are low-extraction ratio drugs, simple allometry was found to be more accurate. Three networks with different input/output formats, namely, format 3 (ANN I), format 4 (ANN II), and a combination of topological indices and input/output format 3, were developed. Three hidden nodes were necessary for these networks to achieve a RMS of 0.02 or less. The overall predictive abilities of ANN I and II were equivalent to or slightly better than those of the allometric method (with indiscriminate use of MLP correction). Application of the MLP correction to only drugs that may be classified as "low extraction ratio" did not improve the predictive ability of the allometric procedure (RMS = 0.12). The low prediction error of ANN II suggests that a pattern regarding the application of MLP correction may be derived, to a large extent, from the animal data.

Categorization of drugs according to their elimination pathways (renal, hepatic, etc.,) may improve the predictive ability of the network, however, these data are most often

Table II. Results for Total Clearance (Cl_{tot})[a]

Drug	Observed	Allometry	Allometry MLP corr.	ANN I	ANN II	ANN III
DDC	0.38	0.49	0.23	0.64	0.37	0.35
SM-1652	0.02	0.04	0.02	0.02	0.02	0.06
Phencyclidine	0.29	1.64	0.74	0.74	0.22	0.10
FCE-22101	0.48	0.24	0.11	0.17	0.24	0.18
Cefotetan	0.04	0.04	0.03	0.14	0.10	0.11
Erythromycin	0.42	1.23	0.34	0.73	1.13	0.54
Methotrexate	0.19	0.17	0.08	0.09	0.13	0.16
Cyclophosphamide	0.20	0.74	0.34	0.68	0.48	0.31
Ceftizoxime	0.21	0.10	0.05	0.82	0.21	0.28
Caffeine	0.12	0.11	0.04	0.23	0.12	0.34
Propafenone	1.09	1.79	0.36	0.55	0.59	1.12
AZT	1.60	1.09	0.49	0.60	0.72	1.15
CAS-94457-09-7	0.28	0.23	0.07	0.21	0.25	0.29
Coumarin	1.40	1.93	0.98	1.17	1.59	1.60
RMS		0.14	0.11	0.11	0.09	0.05

[a] Cl_{tot} expressed as L/hr/kg. ANN I corresponds to input/output format 3. ANN II corresponds to input/output format 4. ANN III corresponds to input/output format 3 with inclusion of topological indices. RMS is the root mean square error.

unavailable for all species. Physicochemical properties of drugs play a very important role in both distribution and elimination. Therefore, the use of structural descriptors was also investigated. Since the predictive abilities of ANN I and II were not dramatically different, input/output format 3 was used along with the structural descriptors to keep the number of input parameters to a minimum. The combination of $^{1}\chi_{v}$, $^{2}\kappa_{\alpha}$, and the animal data format 3 resulted in an overall prediction error of 0.05.

CONCLUSION

This study demonstrates the feasibility of ANNs for multidimensional interspecies scaling of PPs. The various ANNs developed were able to provide adequate generalization even though they were presented with only a limited number of examples. This study also demonstrated the importance of the chemical structure as a supplement to the animal data. However, several limitations of this technique were apparent, the major limitation being the availability of animal data in at least four different species (reduction of the number of species to three or fewer resulted in poor predictions for this data set). Other limitations deal with the development of the ANN itself. For example, there are no set rules and guidelines for selecting the number of hidden nodes, number of training iterations, and preprocessing of data. For this approach to have practical value, the number of animal species will have to be reduced to at least two most commonly used species (for example, rat and dog) and the number of drugs will have to be increased significantly. Future communications will deal with this aspect.

REFERENCES

1. H. Boxenbaum. Interspecies pharmacokinetic scaling and the evolutionary-comparative paradigm. *Drug. Metab. Rev.* **15**: 1071–1121 (1984).
2. H. Boxenbaum. Interspecies scaling, allometry, physiologic time, and the ground plan for pharmacokinetics. *J. Pharmacokin. Biopharm.* **10**:201–227 (1982).
3. R. M. J. Ings. Interspecies scaling and comparison in drug development and toxicokinetics. *Xenobiotica* **20**:1201–1231 (1990).
4. J. Mordenti. Man versus beast: Pharmacokinetic scaling in mammals. *J. Pharm. Sci.* **75**:1028–1039 (1986).
5. H. Boxenbaum. Time concepts in physics, biology, and pharmacokinetics. *J. Pharm. Sci.* 75:1053–1062 (1986).
6. F. E. Yates and P. N. Kugler. Similarity principles and intrinsic geometries: Contrasting approaches to interspecies scaling. *J. Pharm. Sci.* 75:1019–1027 (1986).
7. R. L. Dedrick, K. B. Bischoff, and D. Z. Zaharko. Interspecies correlation of plasma concentration history of methotrexate (NSC-740). *Cancer Chemother. Rep. Part I* 54:95–101 (1970).
8. M. Bonati, R. Latini, G. Tognoni, J. F. Young, and S. Garattini. Interspecies comparison of *in vivo* caffeine pharmacokinetics in man, monkey, rabbit, rat, and mouse. *Drug Metab. Rev.* **15**:1355–1383 (1985).
9. G. S. Duthu. Interspecies correlation of the pharmacokinetics of erythromycin, oleandomycin, and tylosin. *J. Pharm. Sci.* **74**:943–946 (1985).
10. S. S. Ibrahim and F. D. Boudinot. Pharmacokinetics of 2′,3′-dideoxycytidine in rats: Application to interspecies scale-up. *J. Pharm. Pharmacol.* **41**:829–834 (1989).
11. S. M. Owens, W. C. Hardwick, and D. Blackall. Phencyclidine pharmacokinetic scaling among species. *J. Pharmacol. Exp. Ther.* **242**:96–101 (1987).
12. C. Efthymimopoulos, R. Battaglia, and M. Benedetti. Animal pharmacokinetics and interspecies scaling of FCE 22101 a penem antibiotic. *J. Antimicrob. Chemother.* **27**:517–526 (1991).
13. A. Puigdemont, R. Guitart, F. De Mora, and M. Arboix. Prediction of the disposition of propafenone in humans and dogs from pharmacokinetic parameters in other animal species. *J. Pharm. Sci.* **80**:1106–1109 (1991).
14. B. A. Patel, F. D. Boudinot, R. F. Schinazi, J. M. Gallo, and C. K. Chu. Comparative pharmacokinetics and interspecies scaling of 3′-azido-3′-deoxythymidine (AZT) in several mammalian species. *J. Pharmacobio-Dyn.* **13**:206–211 (1990).
15. W. A. Ritschel, N. N. Vachharajani, R. D. Johnson, and A. S. Hussain. Interspecies scaling of coumarin among six different mammalian species. *Meth. Find. Clin. Exp. Pharmacol.* **13**:697–702 (1991).
16. J. Gasfeiger and J. Zupan: Neural networks: A new method for solving chemical problems or just a passing phase. *Anal. Chem. Acta* **248**:1–30 (1991).
17. S. Shadmehr and D. Z. D'Argenio. A neural network for nonlinear Bayesian estimation in drug therapy. *Neur. Comp.* **2**:216–225 (1990).
18. A. S. Hussain, X. Yu, and R. D. Johnson. Application of neural computing in pharmaceutical product development. *Pharm. Res.* **8**:1248–1252 (1991).
19. H. Matsui, K. Yano, and T. Okuda. Pharmacokinetics of the cephalosporin SM-1652 in mice, rats, rabbits, dogs and rhesus monkeys. *Antimicrob. Agents Chemother.* **22**:213–219 (1982).
20. M. Komoya, Y. Kikuchi, A. Tachibana, and K. Yano. Pharmacokinetics of a new broad-spectrum cephamycin YM09330, parentally administered to various experimental animals. *Antimicrob. Agents Chemother.* **20**:176–183 (1981).
21. R. A. Yates, H. K. Adam, R. J. Donnelly, H. L. Houghton, E. A. Charlesworth, and E. A. Laws. Pharmacokinetics and tolerance of single i.v. doses of cefotetan disodium in male caucasian volunteers. *J. Antimicrob. Chemother.* **11** (Suppl A):185–198 (1982).
22. L. B. Mellet. Comparative drug metabolism. *Prog. Drug. Res.* **13**:136–169 (1969).
23. T. Murakama, H. Sakamoto, S. Fukada, S. Nakamoto, T. Hirose, N. Itoh, and M. Nishida. Pharmacokinetics of cefotoxime in animals after parenteral dosing. *Antimicrob. Agents Chemother.* **17**:157–164 (1980).
24. H. C. Neu. The new β-lactamase stable cephalosporins. *Ann. Intern. Med.* 97:408–419 (1982).
25. M. Komuro, W. Hori, M. Hotta, S. Saitoh, R. Ishida, and H. Uchida. Pharmacokinetics of the new antiplatelet agent 2-methyl-3-(1,4,5,6-tetrahydronicotinoyl)pyrazolo[1,5-a]pyridine in laboratory animals. *Arzneim. Forsch./Drug Res.* **42**:55–59 (1992).
26. M. Komuro, W. Hori, M. Hotta, S. Saitoh, R. Ishida, and H. Uchida. Pharmacokinetics of the new antiplatelet agent 2-methyl-3-(1,4,5,6-tetrahydronicotinoyl)pyrazolo[1,5-a]pyridine in human subjects. *Arzneim. Forsch./Drug Res.* **42**:60–64 (1992).
27. L. H. Hall and L. B. Kier. The molecular connectivity Chi indexes and Kappa shape indexes in structure property modeling. In K. B. Lipkowitz and D. B. Boyd (eds.), *Reviews in Computational Chemistry II*, VCH, New York, 1991, pp. 367–422.
28. A. A. Minai and R. D. Williams. Acceleration of backpropagation through learning rate and momentum adaptation. *Int. Joint Conf. Neur. Networks* 1:676–679 (1990).
29. A. S. Hussain and R. D. Johnson. Development of a neural network for QSAR problems. Presented at the 203rd ACS National Meeting (Division of Computers in Chemistry), San Francisco, CA, April 5–10, 1992.
30. R. S. Markin, W. J. Murray, and H. Boxenbaum. Quantitative structure–activity study on human pharmacokinetic parameters of benzodiazepines using the graph theory approach. *Pharm. Res.* 5:201–208 (1988).