QSAR Modeling Using Automatically Updating Correction Libraries: Application to a Human Plasma Protein Binding Model

Sarah L. Rodgers,* Andrew M. Davis, and Nick P. Tomkinson

AstraZeneca R&D Charnwood, Bakewell Road, Loughborough, Leicestershire LE11 5RH, United Kingdom

Han van de Waterbeemd

AstraZeneca R&D Alderley Park, Macclesfield, Cheshire SK10 4TG, United Kingdom

Received June 8, 2007

It is assumed that compounds occupying the same region of model space will be subject to similar errors in prediction, and hence, where these errors are known, they can be applied to predictions. Thus, any available measured data can be used to refine predictions of query compounds. This study describes the application of a correction library to a human plasma protein binding model. Compounds that have been measured since the model was built are entered into the library to improve predictions of current compounds. Time-series simulations were conducted to measure the time dependence of the correction library. This study demonstrates significant improvements in predictions where a library is applied, compared with both a static model and an updating model that includes recently measured data.

1. INTRODUCTION

QSAR models are limited by the data that are used to train them; test compounds outside the compound space representing the training data are more likely to be poorly predicted.[1,2] It has been shown that including the most up to date measurements in model training sets improves their predictive ability.[3] However, it is time-consuming to rebuild models on a frequent basis to account for the large number of measurements made each month. A solution to this problem is to develop a process that automatically updates models as new data are produced.[4] However, systems such as this are at the development stage, and the full benefits are not yet known. An alternative is to have local models generated "on the fly" as predictions are requested.[5,6] This approach involves building local models using only the database compounds considered similar to the query compound. This method is limited by the number of similar neighbors a given query will have; hence, at the beginning of a project, it may not be possible to build a local model due to a lack of neighbors. Automatically updating QSAR models, or automatically generated local models, can be unpopular with scientists who see such "black box" models as difficult to interpret.[7]

A simpler solution is to use the known errors in prediction of compounds similar to the query compound from databases of measurements to refine the predictions of query compounds. These databases of compounds and their associated measured values are known as associative, or correction, libraries[8,9] (hereafter referred to as correction libraries).

The known errors in prediction of existing compounds with measured data are used to make corrections for predictions of similar compounds. This allows the information that can be provided from recently measured compounds to be used without the need for changing the model. Predictions made from global models are able to benefit from "local knowledge" by examining the predicted values of similar neighbors. This method requires far fewer neighbors than would be required to build a local model. In addition, a correction is only made when the neighbors are considered similar enough; thus, corrections are only applied where they are likely to be relevant.

The correction method is based on the similar property principle,[10] which states that compounds with similar structures will also have similar properties. This neighborhood behavior[11] is exploited by the correction libraries approach, which assumes that compounds with similar properties will be subject to the same bias in a model. Hence, the errors in prediction of compounds for which both measured and predicted values are available are used to correct the predictions of similar compounds, for which no measured data are available. By applying the correction, bias is removed from the prediction. All compounds for which measured and predicted values are available can be included in the library.

It has previously been demonstrated that correction libraries provide real benefits in the accuracy of predictions from QSAR models.[12,13] Therefore, we have tested the use of correction libraries on one of our in-house models. We examine the time dependency of the correction libraries, testing how current the libraries need to be in order to be effective.

Human plasma protein binding has been selected as the example model on which to test the performance of correction libraries with time. This is a very important property in drug discovery, and hence there is a lot of in-house measured data and many existing models that attempt to predict it. Although this study is restricted to human plasma protein

* Corresponding author phone: +44(0)1509644436; fax: +44(0)-1509644576; e-mail: Sarah.Rodgers@AstraZeneca.com.

10.1021/ci700197x CCC: $37.00
Published on Web 09/22/2007

binding, we anticipate that the findings of this research will be applicable to many other global models.

2. METHODS

Human plasma protein binding measures the reversible association of drugs with serum albumin and other plasma proteins in the blood. The affinity of a given drug for the various plasma proteins, and the total capacity of these proteins, controls the equilibrium between protein-bound drugs and the free fraction in the plasma. It is thought that only free drug is available to elicit the pharmacological response;[14] hence, plasma protein binding affinity is a key property for potential drug candidates. In addition, it influences many other factors such as the extent of distribution of the drug throughout the body, the rate of metabolism, and renal excretion.

The extent of plasma protein binding can be determined using equilibrium dialysis, as has been previously described.[15] In brief, a plasma-containing equilibrium dialysis cell spiked with a test compound is separated from a cell containing a buffer solution by a semipermeable membrane. Unbound test compound in the plasma-containing cell is able to cross the semipermeable membrane (the proteins are not), and an equilibrium is reached. The amount of test compound in the buffer solution provides an indication of the amount of bound and free test compound in the plasma cell. The dependent variable to be modeled, log K, is a pseudo-equilibrium constant, calculated from the fraction of bound and free drug:

$$\log K = \log \%\ \text{bound} - \log \%\ \text{free} \tag{1}$$

The quality of QSAR models is measured using a range of statistics, providing guidance on how effectively a model fits the data and confidence in how the model will perform in making predictions. External test sets, ideally, temporal test sets, are often employed to test a model's performance, providing a critical measure of the predictivity. The root-mean-square error in prediction (RMSE) reflects deviations of the predicted from the observed value, and the mean error (ME) indicates whether there is any bias in the predictions; they are calculated as follows:

$$\text{RMSE} = \sqrt{\frac{\sum_{i=1}^{n} (y_i - \hat{y}_i)^2}{n}} \tag{2}$$

where n is the total number of compounds, $\hat{y}_i$ is the predicted

$$\text{ME} = \frac{\sum_{i=1}^{n} (y_i - \hat{y}_i)}{n} \tag{3}$$

dependent value, and y_i is the observed dependent value.

The work described here follows the approach of Bruneau and McElroy,[16] identifying nearest neighbors by calculating the Mahalanobis distances (MD_{ij}) between the query, i, and library compounds, j, according to eq 4:

$$\text{MD}_{ij} = \sqrt{(x_i - x_j)^{\mathrm{T}} \mathbf{S}^{-1} (x_i - x_j)} \tag{4}$$

where $(x_i - x_j)$ is the column vector of the descriptor value differences between the query x_i and library x_j compounds, $(x_i - \mathrm{x}_j)^{\mathrm{T}}$ is the transposed vector of $(x_i - x_j)$, and $\mathbf{S^{-1}}$ is the inverse covariance matrix of the descriptors x_j of the library compounds. This distance measure is based on the descriptors that are included in the global model, hence, descriptors that should be relevant for the property that is being modeled. The three nearest neighbors for the query compound are identified from the library. A correction is obtained using the following equation:

$$\bar{z}'_i = \hat{y}_i + \frac{\sum_j (y_j - \hat{y}_j)\, F(\xi_{ij})}{\sum_j F(\xi_{ij})} \tag{5}$$

where $\bar{z}'$ is the corrected prediction of compound i, $\hat{y}_i$ is the initial prediction, $(y_j - \hat{y}_i)$ is the error in prediction for the nearest neighbor j, $F(\xi_{ij})$ is a function that evaluates the distance between compound i and its neighbors, j, and $\sum_j F(\xi_{ij})$ serves as a normalizing factor. The degree to which the calculated error in prediction is applied to the query compound depends on the degree of similarity between the query i and neighbor j; this is adjusted by $F(\xi_{ij})$. The more similar compounds i and j are, the greater the degree of correction up to a maximum full correction, $F(\xi_{ij}) = 1$. Alternatively, where compounds i and j are too dissimilar (more than a threshold Mahalanobis distance), no correction is made to the initial prediction, $F(\xi_{ij}) = 0$.

The degree of correction, controlled by $F(\xi_{ij})$, is calculated as follows:

$$\begin{bmatrix} \text{if } \text{MD}_{ij} \le \text{pl, then } F(\xi_{ij}) = 1 \\ \text{else } F(\xi_{ij}) = e \dfrac{-(\text{MD}_{ij} - \text{pl})^2}{2\sigma^2} \end{bmatrix} \tag{6}$$

where pl is a plateau defining the threshold to which a full correction is made (here, pl = 0.75). The correction factor decreases to zero according to a Gaussian function defined by the two parameters e and σ (in this study, $e = 2.72$ and $\sigma = 0.75$). In our method, the Mahalanobis distance is scaled to account for the number of descriptors in the model. Hence, the Mahalanobis distance used in eq 6 is actually the scaled version according to the formula

$$\text{Scaled MD} = \sqrt{\text{MD}_{ij} \times \frac{15}{d}} \tag{7}$$

where d is the number of descriptors included in the model.

In order to test the time dependency of the correction libraries, some simulation studies were set up to monitor the behavior of correction libraries over time. All experimental data measured up to the end of January 2004 formed the first correction library, a total of 3279 compounds (correction library 1). The first test set was created from all experimental data from February 2004, test set 1. Thus, the library used is from the month before the test set. To form the second correction library, test set 1 was added to correction library 1. Test set 2 contained the measured data from March 2004. In total, 20 test sets were created from February 2004 to September 2005 (test set sizes are reported in Table 1); in

Table 1. Monthly Test Sets Summary Results[a]

test set	n	RMSE uncorrected	RMSE corrected	mean error uncorrected	mean error corrected	standard deviation
1	13	0.72	0.66	0.097	−0.01	0.89
2	32	0.59	0.50	0.15	0.12	0.70
3	37	0.50	0.49	0.12	0.20	0.87
4	13	0.68	0.68	0.04	0.12	0.91
5	11	0.47	0.63	0.06	0.05	0.67
6	18	0.53	0.44	−0.02	−0.03	0.91
7	24	0.65	0.66	0.04	0.07	0.60
8	16	0.54	0.59	−0.27	−0.15	0.49
9	30	0.60	0.49	0.04	−0.02	0.68
10	17	0.61	0.53	0.04	0.24	0.91
11	25	0.56	0.48	−0.13	0.06	0.92
12	25	0.70	0.60	−0.05	−0.10	0.66
13	26	0.53	0.51	0.08	0.02	0.82
14	35	0.55	0.54	0.03	0.04	0.55
15	29	0.69	0.64	0.07	0.02	0.74
16	25	0.59	0.48	0.12	0.10	0.60
17	23	0.69	0.56	0.08	0.11	0.72
18	83	0.62	0.50	−0.05	0.01	0.93
19	31	0.73	0.63	0.00	−0.03	0.77
20	25	0.57	0.56	0.05	0.21	0.68

[a] Units for RMSE, mean error, and standard deviation are log K.

each case, all measured data up to the end of the previous month were placed into a corresponding correction library.

Predictions were made using an in-house global model in the presence and absence of the relevant correction library. This was a PLS model built using SIMCA;[17] the training set consisted of experimental data from before the time series began (i.e., pre January 2004), a total of 3191 compounds. The descriptors used in the model, and thus in the correction libraries, were a subset from an in-house set of descriptors. These include topological (2D), geometrical (3D), and electronic (charge-dependent) descriptors. The initial PLS model was constructed using all descriptors; the variable importance was then used to discard many of the descriptors, and a new model built using a reduced set of 20.

Each test set is temporal to both the training set and the library compounds and hence provides effective validation of the predictivity of the methods tested. Three of the test sets have been made available and are provided in the Supporting Information, the initial test set from February 2004, the test set from December 2004, and the final test set from September 2005. For each of these test set compounds, the neighbors from the correction library are also provided in the Supporting Information.

In addition to the monthly test sets, a final temporal test set was created using data from the three months following the test period (October−December 2005), containing 512 compounds. Predictions were made using the global model alone and then in the presence of each of the 21 monthly correction libraries. This simulates the use of the libraries over a longer time period; each month, the correction library is getting closer in time to the temporal set compounds.

A subset of the final temporal test set was selected which could be made public (hereafter referred to as the public temporal subset); the data are provided as Supporting Information. This subset comprises 53 randomly selected compounds, accounting for 10% of the full temporal set. The three nearest neighbors of each subset test compound are also provided in the Supporting Information.

Table 2. RMSE (log K Units) Values of Final Temporal Test Set and Public Temporal Subset with Updating Libraries and Updating Model

	updating correction library		updating model
library/training set	public temporal subset	final temporal test set	final temporal test set
0	0.59	0.60	0.60
1	0.58	0.60	0.60
2	0.58	0.60	0.60
3	0.58	0.60	0.61
4	0.58	0.60	0.61
5	0.58	0.60	0.61
6	0.58	0.60	0.62
7	0.58	0.60	0.61
8	0.58	0.60	0.62
9	0.58	0.60	0.61
10	0.58	0.60	0.62
11	0.58	0.60	0.61
12	0.57	0.60	0.61
13	0.57	0.60	0.61
14	0.56	0.59	0.61
15	0.56	0.58	0.61
16	0.56	0.58	0.60
17	0.55	0.57	0.60
18	0.55	0.57	0.59
19	0.52	0.56	0.59
20	0.48	0.54	0.57
21	0.47	0.53	0.57

3. RESULTS

The numbers of compounds in each monthly test set, in addition to the standard deviation of the log K values, are provided in Table 1. The test sets are relatively small and vary between the months, but this is a typical pattern in a drug company. The RMSEs in prediction with and without a correction library are plotted in Figure 1 and included in Table 1. For the majority of the monthly test sets, the RMSE is reduced when the correction library is used. This reduction in the error of prediction is statistically significant (matched pairs Student t-test, $p < 0.05$) for six of the test sets (2, 9, 12, 16, 18, and 19). For the three test sets where the RMSE of the corrected predictions is greater, generally there are only a small number of compounds present, or alternatively both corrected and uncorrected predictions are poor (RMSE close to standard deviation). The average change in RMSE with the introduction of correction libraries is −0.06 log K units, with a range from −0.16 (a reduction in RMSE with the application of a correction library) to +0.15 (increase in RMSE with the application of a correction library). The latter refers to the test set from June 2004, involving only 11 compounds, and the former to the test set from August 2005, with 31 compounds.

For all three of the published monthly test sets, an improvement in prediction error results when the correction library is applied. For the test set from February 2004, the RMSE reduces from 0.72 to 0.66; for the December 2004 test set, the reduction is from 0.56 to 0.48, and for the final September 2005 test set, it is 0.57 to 0.56. All but one of the compounds from these three test sets receive a correction. The degree of the correction varies according to the Mahalanobis distance of the nearest neighbors and for some compounds can be very small.

Predictions for the final temporal test set were made with the initial in-house model and then with the application of each of the 21 correction libraries from the time series

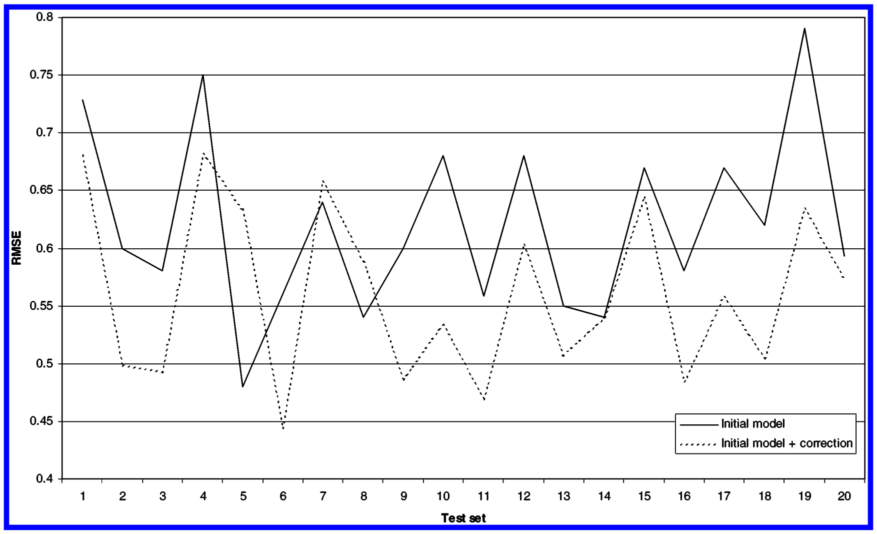

Figure 1. RMSE (log *K* units) values for individual monthly test sets with and without a correction library.

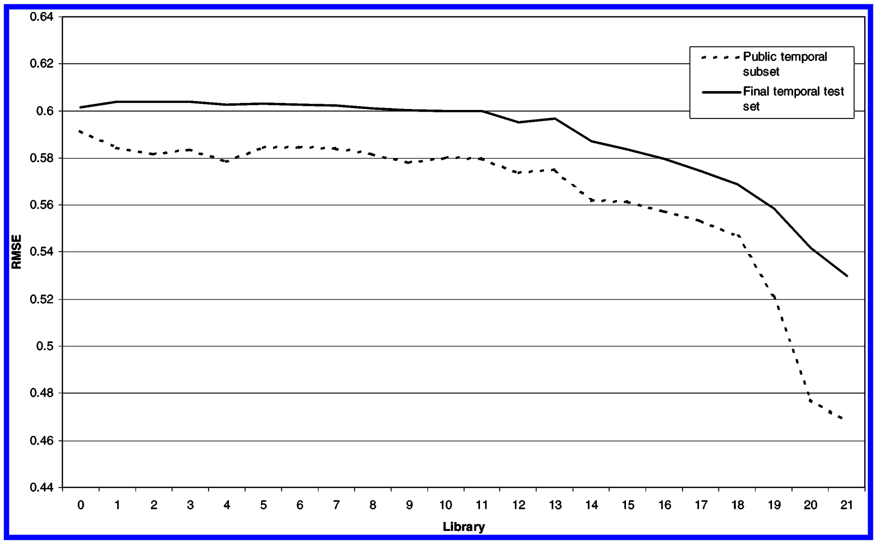

Figure 2. Final temporal test set and public temporal subset prediction errors (log *K* units) with monthly libraries.

(including the library from September 2005 not included in the monthly test set analysis). The error associated with predictions from the library, for the final temporal test set and the public temporal subset, are plotted in Figure 2 (and summarized in Table 2). The first point on this graph, at library point “0”, is the prediction from the model with no correction library; library point 1 shows predictions using the library from January 2004 and so forth. The standard deviation of the public temporal subset is 0.74 log *K* units, and the final temporal test set is 0.81. For approximately the first 10 months, the libraries have no effect on the prediction error of either temporal set. After this period, the RMSE reduces each month as the library is updated, reaching a minimum RMSE of 0.53 in the final month for the final temporal test set (initial RMSE = 0.60). For the public temporal subset, there is an even greater reduction, with the RMSE decreasing from an initial value of 0.59 to 0.47 with the final library. The mean error associated with the corrected predictions for the final temporal test set is low and remains stable across the time series (ranging from −0.052 to −0.028). The mean error for the public temporal subset is much greater, ranging from −0.11 to −0.17, but again remains stable across the time series. The larger error associated with the public temporal subset is likely to be a product of the small number of compounds considered.

The reduction in the error in prediction for the final temporal test set when the libraries are applied starts to become statistically significant (matched pairs Student *t*-test, $p < 0.05$) at library 14 (compounds up to February 2005). The statistical significance of the reduction of error in prediction improves the following month to $p < 0.005$ and then again with library 17 (compounds up to May 2005) to $p < 0.001$. Hence, the reduction in the errors in prediction from the final eight libraries when compared with no library are statistically significant, with the level of significance increasing over time. For the public temporal subset, the reduction in error in prediction is only significant for the final two libraries (matched pairs Student *t*-test $p < 0.05$).

The RMSE values reported above include those compounds that do not receive a correction when the libraries are applied, and hence the predictions (and error in predictions) do not change. If only those compounds that receive a correction are considered, the RMSE of the final temporal test set reduces from 0.60 to 0.51 when the final library is applied. For the public temporal subset, the RMSE of corrected compounds reduces from 0.55 with no library to 0.39 with the final library.

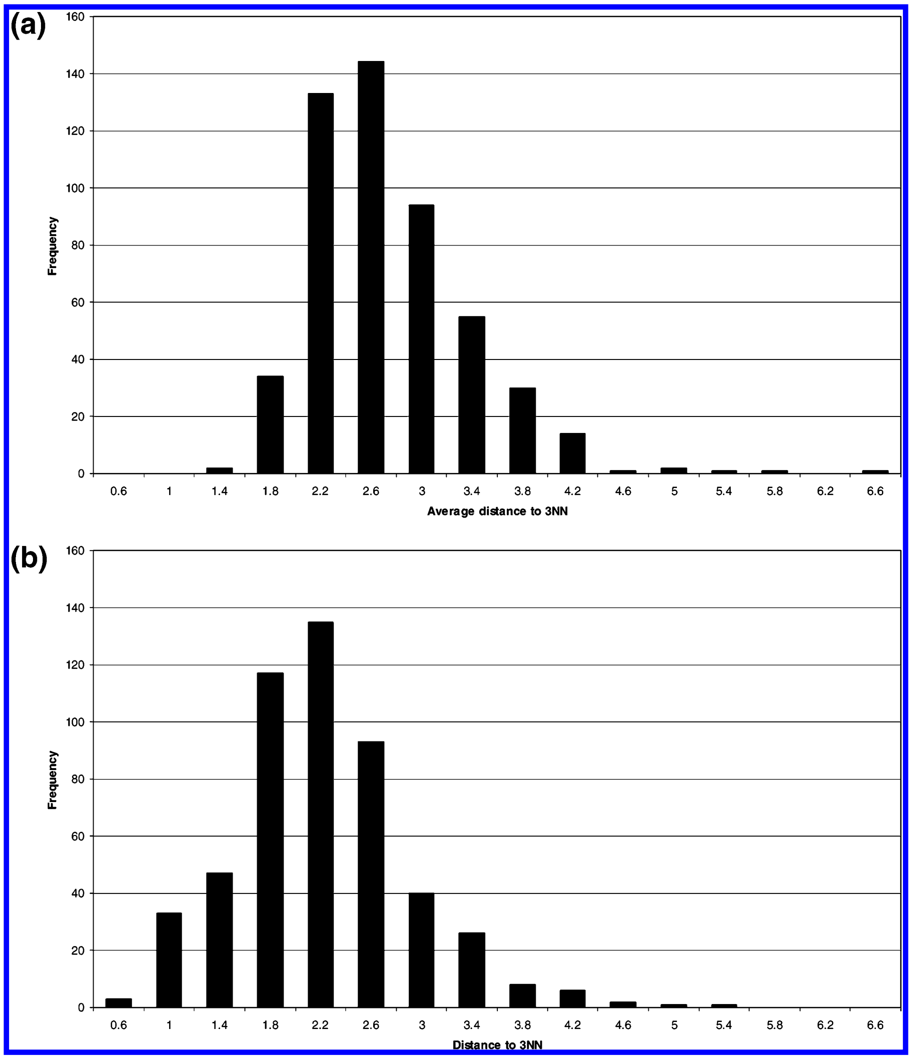

Figure 3. Frequency distribution of distance to the nearest neighbors of final temporal test set compounds to (a) the first correction library (data to January 2004) and (b) the final correction library (data to September 2005).

For the public temporal subset, 31 compounds receive a correction from the first library of less than 0.1 log *K* units (from a total of 53 compounds), including three compounds that received no correction. With the final library, two received no correction and 16 received a correction of less than 0.1. This pattern is reflected with the final temporal test set: the number of compounds receiving no correction reduces from 146 compounds with the first library to 52 compounds with the final library; 472 compounds receive a correction of less than 0.1 log units with the first library; this decreases to 285 compounds when the final library is applied.

As more compounds are added to the library, there is an increased probability of nearest neighbors being identified for each query compound, and thus the number of corrections increases. As the library gets closer in time to the temporal test set, the new compounds added to the library each month are likely to be more similar to the final temporal test set compounds. This is not the case with the public temporal subset; the average Mahalanobis distance to the three nearest neighbors does not reduce between the initial and final libraries. However, with the final temporal test set, there is a significant reduction in the average distance to the three nearest neighbors over time with an average Mahalanobis distance of 2.54 with the first library and 2.01 with the final library. This corresponds to a RMSE in prediction of 0.60 when the first library was applied, reducing to 0.53 when the final library was applied. The frequency distributions of these distances for the final temporal test set compounds are plotted in Figure 3.

4. DISCUSSION

Correction libraries provide an effective tool of "localizing" global models. By identifying three nearest neighbors, the prediction refinement process is local in nature. The time period over which data for a library was collected has a clear effect on the accuracy of the predictions. Old libraries will provide little benefit in making corrections for current compounds; however, old libraries will also cause no deterioration of the predictions. Figure 2 clearly displays this relationship with time. Library compounds measured more than approximately 10 months before the current (query) compounds are not able to positively influence predictions. After this 10-month threshold, as the time of measurement of the library compounds gets closer to that of the query compounds, the predictions improve.

A previous publication[3] examined the automatic updating of a QSAR model on a regular basis to reflect more recent experimental measurements. The same data added to the

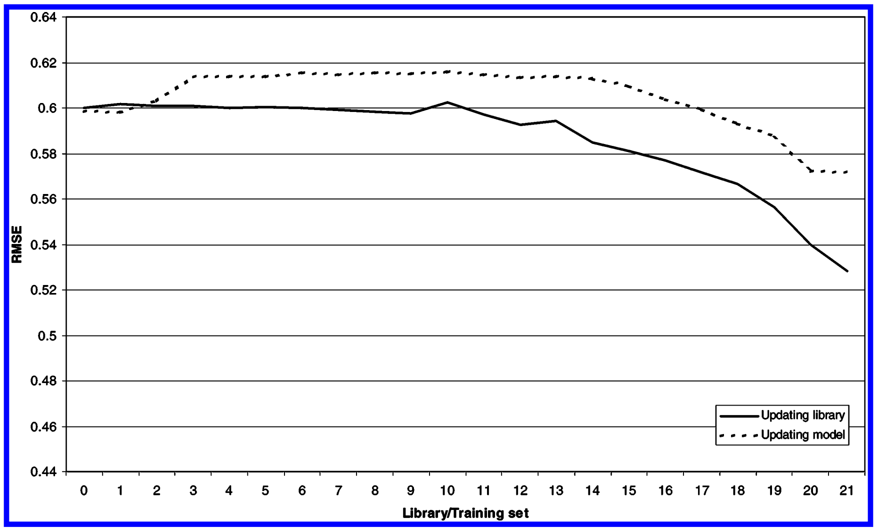

Figure 4. Final temporal test set prediction error (log *K* units) with time-series libraries (standard deviation = 0.81).

libraries each month in the time-series analysis described here were used to update a training set for the same model. The model was updated using the same descriptor set used for the initial in-house model (and thus for the correction library calculations used here). Hence, the two methods of updating predictions can be directly compared (the final temporal test set used here is slightly smaller than that in the previous study, only the compounds considered here are included in the subsequent analysis).

The RMSE values resulting from predictions made using the series of models built for human plasma protein binding are presented in Figure 4 (and summarized in Table 2). The errors associated with the predictions for the final temporal test set predictions where a correction library is applied are included for comparison. Each month, the newly measured compounds are either included in the training set and a new model is built or they are added to the library. For both the updated model and the updated library, the errors in the predictions reduce over time; however, the updating library offers a greater reduction in error than the updating model. The updating models do not improve (in terms of RMSE compared with initial model) until model 18 (compounds to June 2005), with only the final four models of the time series performing better than the initial model. Model 19 provides a statistically significant (matched pairs Student *t*-test, $p < 0.05$) reduction in prediction error as compared to model 1. This is also the case for models 20 ($p < 0.01$) and 22 ($p < 0.005$). As detailed in the Results section, the libraries offer a statistically significant reduction in prediction error from library 14 onward.

It would appear that the recently measured compounds are able to exert more influence on predictions when included in a correction library than they are when added to a training set. The same compounds were used in the updating libraries and updating training sets; hence, the same information was made available in the two methods. The measured data in the library are directly applied to query compounds for which neighbors exist; alternatively measured data in the training set is part of a large volume of data used to build a global model.

Including newly measured compounds in a correction library is a much simpler approach than rebuilding the model to include those compounds in the training set. It is also likely to be more acceptable as it does not involve changing the QSAR model. Since only three nearest neighbors are used to correct a prediction, there is a high probability of identifying neighbors for each query compound. However, the effectiveness of the correction libraries over a longer time period is still unknown and warrants further analysis. Eventually, it is likely that it will become necessary to update models with more recent measurements in the training set.

The neighbors selected to provide corrections are not always from the same project as the query compound. For the final temporal test set used in the time-series analysis, 32% were corrected by three neighbors from the same project. For 41%, two of the neighbors belonged to the same project as the query, and 50% had at least one neighbor from the same project. This may appear surprising, but it highlights the similarities in the properties that determine the QSAR model that exist between the chemical series used in different projects. Global models exploit these similarities to produce a general predictive tool for the property being modeled.

The success of the correction libraries in improving predictions suggests that there may be little or no requirement for a QSAR model. Predictions can be calculated from the experimental values of the three nearest neighbors of each query compound. This proposition has been tested; the final temporal test set provided the set of query compounds and the final library the compounds from which the nearest neighbors were identified. Three different methods for the identification of the nearest neighbors were tested: (1) using the set of 20 descriptors used in the human plasma protein binding models (and therefore identified as being important for plasma protein binding) and calculating the Mahalanobis distance, (2) using all available descriptors from an in-house set (total = 194) and calculating the Mahalanobis distance, and (3) using in-house fingerprints similar to the Daylight fingerprints and the Tanimoto similarity coefficient. For methods 1 and 2, the Mahalanobis distance threshold used by the correction libraries was applied; for method 3, only compounds with a Tanimoto similarity > 0.7 were considered as neighbors.

The results are summarized as the RMSE values of the compounds for which nearest neighbors could be identified in Table 3. Also included are the RMSE values for the initial model, the initial model with the application of a library,

Table 3. RMSE in Prediction (log *K* Units) of Updating Models and Correction Libraries with a Naïve Nearest-Neighbor-Based Prediction

	initial model	initial model + correction library	updated model	nearest-neighbor model	*N*
model descriptors	0.59	0.48	0.57	0.58	334
all descriptors	0.60	0.50	0.58	0.58	360
fingerprints	0.66	0.50	0.63	0.54	183

and an updated model. In each case, nearest neighbors could not be identified for all of the test compounds; for the fingerprint-based neighbors, less than half of the test compounds were predicted. For the two descriptor methods, around 65% of the compounds received a correction. Hence, such an approach is not practical, as predictions cannot be obtained for a large proportion of the test set.

However, the naïve nearest-neighbor predictor performs surprisingly well considering the simplicity of the approach; the errors in prediction for each test subset (= compounds that have nearest neighbors) are comparable to that of our descriptor-based QSAR model. This demonstrates how powerful predictions based on measurements of other compounds can be and hence suggests why the correction libraries are so successful. However, in each case, using both the descriptor-based QSAR model and a correction library provides the most accurate predictions. The combination of these two methods presents the best approach for making accurate predictions.

Correction libraries may be applied to any QSAR model for which the data to build a library are available. This study examined the use of correction libraries on a human plasma protein binding PLS model. However, correction libraries are not limited to PLS models and may be applied to predictions calculated from any QSAR model. Future studies are planned to examine the application of correction libraries to other QSAR models including local models and models built using methods other than PLS.

5. CONCLUSION

Correction libraries have been shown here, and in other publications,[12,13] to provide a real benefit in improving the predictions of QSAR models. The libraries provide a simple tool for refining predictions and here have shown greater reductions in RMSE values than where the model itself was updated. However, the initial QSAR model is still very important; the best predictions were obtained where a descriptor-based QSAR model and correction library were used.

Supporting Information Available: A series of text files containing the test and library data described. The test sets from February 2004, December 2004, and September 2005 are available, each containing the descriptors used in the model and the measured values. Associated with these is a library file, containing the descriptors, measured and predicted values, and the date the compound entered the library. The public temporal subset is also made available, again with a separate library file. (A total of six separate files.)

REFERENCES AND NOTES

(1) Stouch, T. R.; Kenyon, J. R.; Johnson, S. R.; Chen, X.-Q.; Doweyko, A.; Li, Y. In silico ADME/Tox: why models fail. *J. Comput.-Aided Mol. Des.* **2003**, *17*, 83−92.

(2) Tetko, I. V.; Bruneau, P.; Mewes, H.-W.; Rohrer, D. C.; Poda, G. I. Can we estimate the accuracy of ADME-Tox predictions? *Drug Discovery Today* **2006**, *11*, 700−707.

(3) Rodgers, S. L.; Davis, A. M.; Van de Waterbeemd, H. Time-Series QSAR Analysis of Human Plasma Protein Binding Data. *QSAR Comb. Sci.* **2007**, *26*, 511−521.

(4) Cartmell, J.; Enoch, S.; Krstajic, D.; Leahy, D. E. Automated QSPR through Competitive Workflow. *J. Comput.-Aided Mol. Des.* **2005**, *19*, 821−833.

(5) Guha, R.; Dutta, D.; Jurs, P. C.; Chen, T. Local Lazy Regression: Making Use of the Neighbourhood to Improve QSAR Predictions. *J. Chem. Inf. Model.* **2006**, *46*, 1836−1847.

(6) Zhang, S.; Golbraikh, A.; Oloff, S.; Kohn, H.; Tropsha, A. A Novel Automated Lazy Learning QSAR (ALL-QSAR) Approach: Method Development, Applications, and Virtual Screening of Chemical Databases Using Validated ALL-QSAR Models. *J. Chem. Inf. Model.* **2006**, *46*, 1984−1995.

(7) Gola, J.; Obrezanova, O.; Champness, E.; Segall, M. ADMET Property Prediction: The State of the Art and Current Challenges. *QSAR Comb. Sci.* **2006**, *25*, 1172−1180.

(8) Tetko, I. Neural Network Studies. 4. Introduction to Associative Neural Networks. *J. Chem. Inf. Model.* **2002**, *42*, 717−728.

(9) Tetko, I. Application of Associative Neural Networks for Prediction of Lipophilicity in ALOGPS 2.1 Program. *J. Chem. Inf. Model.* **2002**, *42*, 1136−1145.

(10) Johnson, M. A.; Maggiora, G. M. In *Concepts and Applications of Molecular Similarity*; Wiley: New York, 1990.

(11) Patterson, D. E.; Cramer, R. D.; Ferguson, A. M.; Clark, R. D.; Weinbuerger, L. E. Neighbourhood Behaviour: A Useful Concept for Validation of "Molecular Diversity" Descriptors. *J. Med. Chem.* **1996**, *39*, 3049−3059.

(12) Tetko, I. V.; Bruneau, P. Application of ALOGPS to Predict 1-Octanol/Water Distribution Coefficients, logP, and logD, of AstraZeneca In-House Database. *J. Pharm. Sci.* **2004**, *93*, 3103−3110.

(13) Tetko, I. V.; Tanchuk, V. Y. Application of Associative Neural Networks for Prediction of Lipophilicity in ALOGPS 2.1 Program. *J. Chem. Inf. Model.* **2002**, *42*, 1136−1145.

(14) Van de Waterbeemd, H.; Gifford, E. ADMET in silico modelling: Towards Prediction Paradise? *Nat. Rev. Drug Discovery* **2003**, *2*, 192−204.

(15) Fessey, R. E.; Austin, R. P.; Barton, P.; Davis, A. M.; Wenlock, M. C. The Role of Plasma Protein Binding in Drug Discovery. In *Pharmacokinetic Profiling in Drug Research*, 1st ed.; Testa, B., Krämer, S. D., Wunderlie-Allenspach, H., Folkers G., Eds.; Verlag Helvetica Chimica Acta: Zurich, Switzerland, 2006; pp 119−141.

(16) Bruneau, P.; McElroy, N. R. *logD*$_{7.4}$ Modelling Using Bayesian Regularized Neural Networks. Assessment and Correction of the Errors of Prediction. *J. Chem. Inf. Model.* **2006**, *46*, 1379−1387.

(17) *SIMCA-P*, v10.0.4.0; Umetrics: Umeå, Sweden 2003. Available from http://www.umetrics.com.

CI700197X