Grouping Pharmacokinetic Profiles Using Kohonen Self-Organizing Maps

Jason L. Boyd, Robert Fraczkiewicz, Grazyna Fraczkiewicz, Robert D. Clark, and Walter S. Woltosz
Simulations Plus, Inc., 42505 10th Street West, Lancaster, CA 93534, USA (www.simulations-plus.com)

INTRODUCTION

The shapes of plasma concentration versus time (Cp-time) profiles from large clinical trials are often highly variable, even in well-controlled trials involving homogeneous cohorts. Sources of variability include genetic, developmental, demographic, dietary, and lifestyle differences. Pharmacokinetic scientists are tasked with assessing the variabilities in these data sets and identifying mechanistic sources of those variabilities. Rapidly processing data to identify and graphically depict trends is a standard practice in pharmacokinetic modeling which, when properly implemented, can effectively guide modeling efforts. Kohonen self-organizing maps are an unsupervised clustering tool for organizing heterogeneous data, and this poster demonstrates that they can be a tool that pharmacokinetic scientists may find useful in understanding complex, highly variable clinical trial results.

ABOUT KOHONEN MAPS

Kohonen mapping produces a nonlinear transformation of input descriptors that maps observations onto a two-dimensional grid of cells in such a way that the observations mapping to the same or nearby cells are similar in terms of the original descriptors. Observations mapping to the same or adjacent cells represent a cluster, with cells to which no observations map serving to separate clusters containing dissimilar observations from each other. Fig. 1 shows a chem-informatics application in which the observations are drug molecules and the mapping was based on molecular descriptors. The colored dependent variable was measured *in vitro* apparent permeability, which was *not* used to generate the self-organizing map.

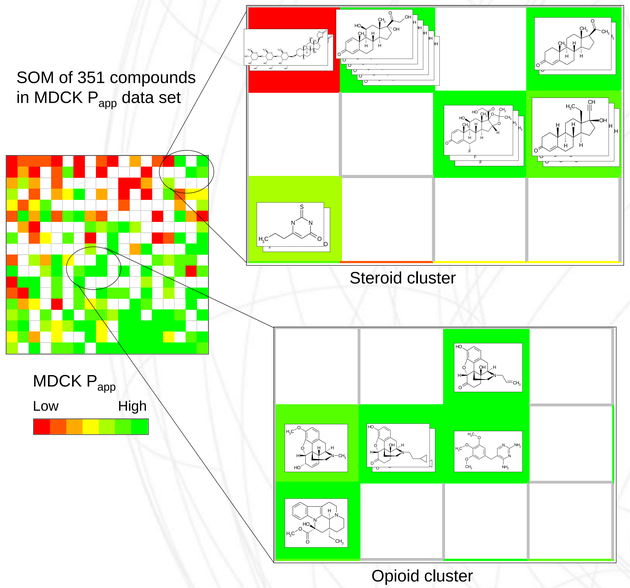

***Figure 1**: Use of Kohonen self-organized mapping to relate molecular structure to apparent permeability (P_{app}) for MDCK (Martin-Darby Canine Kidney) cells. The colored map shows how the 351 compounds in the data set distribute across the map. The cells in the map have been colored according to the average permeability for all compounds in each cell, with green indicating the highest average permeability, red indicating the lowest average permeability, and white indicating empty cells.*

Kohonen T. Self-organized formation of topologically correct feature maps. Biological Cybernetics, 1982, 43: 59 – 69.

DATA

Cp-time profiles were provided by a pharmaceutical client* and come from a clinical trial in 93 consenting volunteer subjects† selected using criteria intended to minimize variability, including a narrow range of age, body weight, and body mass index. Dietary intake, concomitant medications, and physical activity were also constrained. Each subject was administered a 24-hour controlled release formulation containing the same dose of a developmental drug*. The Cp-time profiles obtained varied by three orders of magnitude in peak concentration, and the time required to reach peak concentration ranged from one hour to over 20 hours (Fig 2). This high variability in pharmacokinetics was unrelated to demographic variables (age, body weight, BMI, ethnicity, gender). Moreover, *in vitro* dissolution was not predictive of *in vivo* results. The client was concerned that the observed variability was related to formulation effects such as excipient interactions, or other conditions for which data had not been collected.

* The company providing the data prefers that the compound name, structure, formulation details, and identity of company not be divulged here.

†The trial was approved by an industrial Investigational Review Board.

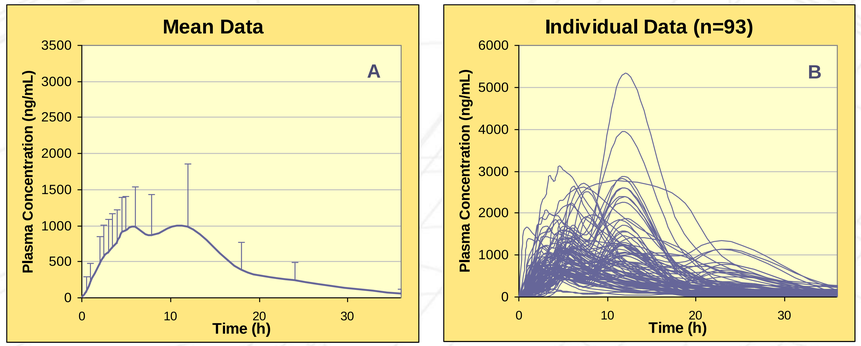

***Figure 2**: A) Mean concentration versus time profile. B) The individual concentration versus time profiles for the 93 subjects in the trial.*

ANALYSIS & RESULTS

The data was organized in such a way that each time point was treated as a separate descriptor and the Cp-time profile for each subject treated as a separate observation (Fig. 3). The default grid size is a square of size equal to the largest squared number that is less than the number of observations; however, 6x6 was used here. Applying Kohonen mapping yielded six groups characterized by shared features such as sharp early peaks, sharp late peaks, rounded late peaks, and profiles with multiple peaks. The smallest cluster contained one profile and the largest cluster contained 27 profiles. The Kohonen map obtained, color-coded by group, is shown in Fig. 4 and the profiles within each group are shown in Fig. 5.

A

	A	B	C	D	E	F	G	H
1	TIME	0	0.5	1	2	2.5	3	3.5
2	S201	0	0	150	499	692	791	1110
3	S202	0	307	641	939	826	712	690
4	S203	0	25.4	254	914	1420	1440	1410
5	S204	0	19.4	425	986	1190	1120	1020
6	S205	0	300	529	442	514	599	658

B

	A	B	C	D	E	F	G
39	NUMBER	NAME	SET	CELL	CLUSTER	ROW	COLUMN
40	1	S201	Test	3	4	0	3
41	2	S202	Train	8	2	1	2
42	3	S203	Test	33	4	5	3
43	4	S204	Train	10	2	1	4
44	5	S205	Train	32	1	5	2

***Figure 3**: Tab-delimited data files involved in Kohonen mapping by ADMET Predictor. A) Input file with plasma concentration profiles. Each row contains data for one subject, each column represents one sampling time, and entries in individual cells are concentrations. B) Output file showing the map cell, row, and column to which each profile was assigned. The "cluster size" is the number of observations assigned to that cell.*

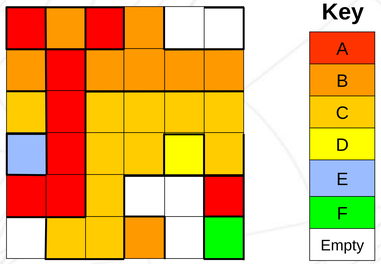

***Figure 4**: Final Kohonen map for the 93 Cp-time profiles used in this study, with each color representing a different cluster. Clusters were generated by combining cells in the Kohonen map which share similar profile shapes but differ in the maximum concentration reached. Boundaries between clusters are indicated by bold lines. Note that the underlying map is toroidal, so cells at the left edge are adjacent to those at the right, and the top is adjacent to the bottom.*

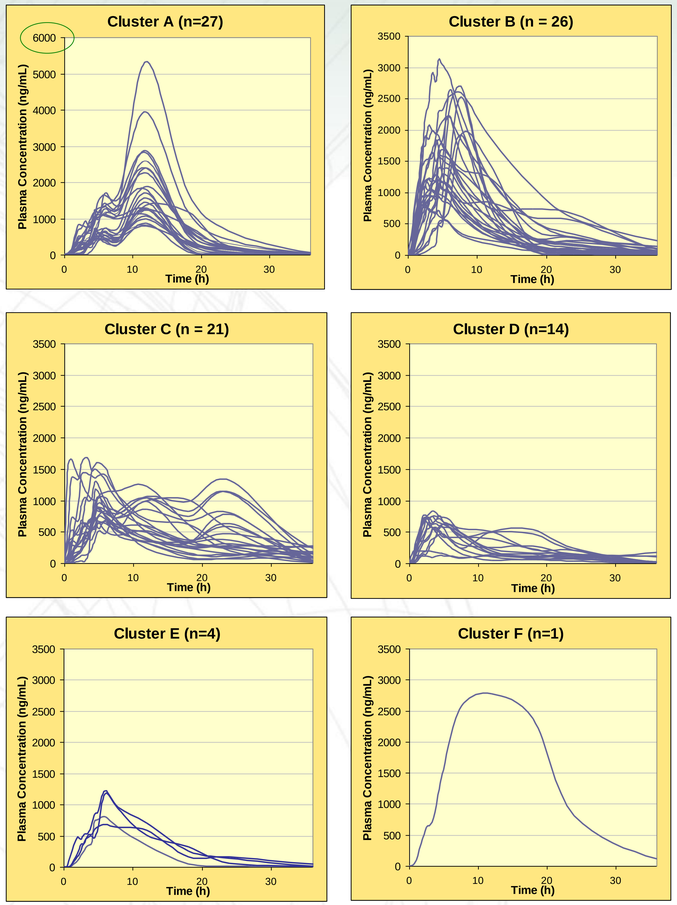

***Figure 5**: Subgroups of Cp-time profiles identified using the Kohonen mapping tool in ADMET Predictor.*

Clearly the mean profile (Fig. 2A) fails to capture the true extent of the variability in individual data (Fig. 2B). Kohonen mapping generated subgroups of Cp-time profiles with more consistent shapes (Figure 5), which proved invaluable in guiding subsequent analysis. Splitting the data set into subgroups made it possible to create a mechanistic pharmacokinetic model (not shown) that identified differences in dissolution, absorption, and metabolism as a way to explain observed data. In particular, differences among the groups was attributed to differences in drug dissolution *in vivo*. Variability within each group, on the other hand, was attributed to interindividual variation in metabolism and variable gastric retention of the tablet.

CONCLUSION

Identifying subgroups in Cp-time profiles by Kohonen mapping can be a useful way to investigate sources of variability in clinical trials where demographic and clinical trial data alone are insufficient to rationalize the observed variabilities. Application of this technique to pharmacokinetic data analysis has the potential to significantly improve the efficiency and productivity of pharmaceutical research.

simulationsplus, inc.