Journal of Scientific Research & Reports
3(22): 2884-2899, 2014; Article no. JSRR.2014.22.006
ISSN: 2320–0227

SCIENCEDOMAIN *international*
www.sciencedomain.org

Determination of Dermal Permeability Coefficient (Kp) by Utilizing Multiple Descriptors in Artificial Neural Network Analysis and Multiple Regression Analysis

Ronald Bartzatt[1*]

[1]*University of Nebraska, Durham Science Center, 6001 Dodge Street, Omaha Nebraska 68182, United States.*

Author's contribution

The sole author designed, analyzed and interpreted and prepared the manuscript.

Article Information

DOI: 10.9734/JSRR/2014/13125
Editor(s):
(1) William Ebomoyi, Department of Health Studies, College of Health Sciences, Chicago State University, USA.
Reviewers:
(1) Rahul Rama Hegde, School of Pharmaceutical Sciences, IFTM University, Lodhipur Rajput, Moradabad 244102, India.
(2) Anonymous, Hawler Medical University, Iraqi Kurdistan, Region.
Peer review History: http://www.sciencedomain.org/review-history.php?iid=637&id=22&aid=6045

Original Research Article

Received 8th August 2014
Accepted 25th August 2014
Published 10th September 2014

ABSTRACT

Aims: The permeability coefficient, or Kp, is an important descriptor for assessing dermal absorption of medicaments utilized for clinical treatment of various dermal accessible diseases. Determination of Kp by multiple descriptors by artificial neural network (ANN) and multiple regression is compared.
Study Design: The calculation of Kp utilizing multiple descriptors, and comparison of ANN and multiple regression is achieved.
Place and Duration of Study: Durham Science Center, Chemistry Department of the University of Nebraska, between April 2014 and July 2014.
Methodology: The calculation of Kp by previous methodologies is accomplished for a

**Corresponding author: E-mail: rbartzatt@unomaha.edu;*

broad spectrum of medicinal and chemical compounds. The values Kp thus acquired are then compared to those obtained by ANN training and multiple regression analysis. Various other pharmaceutical based descriptors are then applied to ascertain the benefit of Kp determination by those properties.

Results: Training and determination of Kp by ANN showed that Log Ko/w and molecular weight (MW) utilized by conventional means is effective. However, ANN demonstrated the Kp determination by applying properties of Log Ko/w, MW, polar surface area, number of atoms, rotatable bonds, molecular volume, and atoms responsible for hydrogen bond donor and acceptors, are also effective and offer significant advantages. These advantages include the potential of encompassing many more molecular constitutional descriptors and molecular properties. Multiple regression showed clearly that the application of more descriptors for Kp determination increases the coefficient of determination (R^2). Increased R^2 shows an improved fit of the raw data to the model improved prediction.

Conclusion: Determination of Kp by applying various descriptors in addition to Log Ko/w and MW increases the model fit to the raw data. ANN prediction of Kp was more effective when using additional descriptors. Prediction of Kp by multiple regression was useful, and utilizing descriptors with Log Ko/w and MW improved the model fit to the raw data.

Keywords: Dermal permeability coefficient; descriptors; dermal; skin sbsorption.

ABBREVIATIONS

Term: Kp, dermal permeability coefficient; PSA, polar surface area; nON, number of oxygen & nitrogen atoms; nrotb, number of rotatable bonds; MW, molecular weight; nOHNH, number of hydroxyl groups & amine groups; MV, molecular volume; microns, μ; um, micrometer; MR, multiple regression; ANN, artificial neural network; Log Ko/w, partition coefficient between 1-octanol and water layers.

1. INTRODUCTION

Transdermal drug delivery has made many important contributions to medical practice. The therapeutic benefits of administering drugs through the skin has been recognized for many years. Key to successful use of topical medicaments is understanding the movement of drugs through the dermal barriers. The structure of the skin is highly organized, heterogeneous, and multilayered [1]. Summation of the various layers that comprise the epidermis and dermis, the appendages, and underlying microvasculature comprise a living envelope that envelopes the body [1]. The permeability coefficient Kp (cm/hour), is a principal parameter in estimating dermal absorption of drugs. The effective use of Kp values in assessing compound crossing of the dermal layers necessitates the understanding of processes involved with drug transport across the skin.

The skin is broadly considered to be composed of two layers [1]: 1) The epidermis, which is a nonvascular layer of approximately 100 *μ*m thick; 2) The highly vascularized dermis of about 500 um to 3,000 μm thick. The layer thought to provide the major barrier to the absorption into the circulation of most substances deposited on the skin surface is the outermost layer of the epidermis, which is the stratum corneum is about 10-40 μm thick. Often measured in fluxes, or mass transport of molecules moving through a cross-sectional area per time, the diffusional flux is effectively described by Fick's first law (postulates that

flux goes from high concentration to low concentration) [2,3]. Fick's second law predicts how diffusion causes the concentration to change with time. Both laws contain the diffusion coefficient (D), which being often difficult to assess, the second law can be abbreviated by introducing the permeability coefficient (Kp) [2,3]. The permeability coefficient makes easier the evaluation of topical drug usage and effectiveness [4].

Advantages of transdermal drug delivery include [2-6]: 1) Avoids degradation that can occur in the stomach environment; 2) Avoids enzyme conversion to inactive molecules in many cases; 3) Provides steady plasma levels; 4) Transdermal delivery systems are easy to use and noninvasive; 5) Increases patient compliance; 6) Deceases medical wastage and reduces cost. The most commonly utilized descriptor to represent the diffusion of compounds is the permeability coefficient (Kp) [7]. Earlier studies have shown that transdermal administration of drugs can also increase the therapeutic index with a simultaneous decrease in side effects [8].

The skin is considered the largest organ in the body and it protects against the influx of toxins, water, and is generally impermeable to the penetration of foreign molecules [9]. The often times observed variations among individuals in the rate at which drugs are absorbed through the skin is due to such factors such as thickness of the stratum corneum, skin hydration, underlying skin diseases/injuries, ethnic differences, and body temperature [9]. Model prediction of Kp by application of quantitative structure activity relationships are in fair agreement with experimental data for those compounds that are applied in a water-based vehicle when the integrity of the skin was not compromised [10].

Almost all the skin's barrier properties is due to the superficial layers of the epidermis, the stratum corneum [11,12]. It is usually viewed as unlikely that noticeable absorption occurs through sweat pores and hair follicles [11]. However, the major modes of diffusion include intercellular, transcellular, and transappendageal [1,11]. A local or even systemic medicinal effects can be achieved by administration of topical drugs [13].

Among the factors affecting the dermal penetration by compounds include the location of the compound on the body, age of individual, extent of skin covered by the compound, and condition of the skin [1,2]. The rate of penetration into the skin can be quantitatively assessed by use of the permeability coefficient (Kp) [1]. Previous studies have determined a highly effective mathematical model for Kp determination, by applying the compound's molecular weight (MW) and octanol-water partition coefficient Log $K_{o/w}$ (Log [1-octanol phase/aqueous phase]) [14-17]:

$$\text{Log } Kp = -2.72 + 0.71(\text{Log } K_{o/w}) - 0.0061(MW) \quad (1)$$

Some studies have shown that the descriptors important for influencing dermal penetration by a compound are consistently hydrophobicity (represented by Log $K_{o/w}$), molecular size (represented by molecular weight or molecular volume), and hydrogen bonding capability (represented by H-bond donors $–NH_n$ and –OH) [14]. In general, very large molecules penetrate the skin slower than small molecules, hence the use of molecular weight to model permeability [15]. Other studies have shown that Kp prediction based on molecular volume is just as effective as values based on molecular weight. The use of equation (1) has been shown to have good predictive capability, producing effective results [1,17].

Therefore, the study and consideration of topical medicaments is advantageous for clinical treatment of numerous disease conditions. The ease, versatility, and accuracy of applying

the permeability coefficient (Kp) will remain important. Exploration of molecular properties that can be applied for characterization and calculation of (Kp) will be beneficial for development and understanding of topical pharmaceuticals.

2. METHODOLOGY

2.1 Molecular Modeling and Physicochemical Properties

The molecular structures of all compounds were visualized by utilizing ACD/Chem Sketch version 12.01 (copyright © 1994-2009, Advanced Chemistry Development, Inc. 8 King Street East, Suite 107, Toronto, Ontario, Canada M5C 1B5) and Chem Windows 3 version 3.0.2 (Softshell International, 715 Horizon Dr. Ste 390, Grand Junction C0 81506 USA). Various properties/descriptors were determined by Molinspiration Cheminformatics (Nova Ulica, SK-900 26 Slovensky Grob, Slovak Republic), Osiris Property Explorer (Actelion Pharmaceuticals Ltd., Gewerbestrasse 16, 4123 Allschwil, Switzerland), and Molsoft property window (copyright Molsoft L.L.C., 11199 Sorrento Valley Road, S209 San Diego CA 92121 USA).

Various descriptors required for the study were determined by Molinspiration Cheminformatics and Osiris Property Explorer. Values of Kp (permeability coefficient) were determined by DERWIN v1.42 (copyright © 2000 U.S. Environmental Protection Agency, Washington D.C. USA). Diffusion coefficient (D) was determined by Physiology Web Diffusion time calculator, last accessed June 15 http://www.physiologyweb.com/calculators/diffusion_time_calculator.html.

2.2 Artificial Neural Network (ANN) Methodology and Software

Artificial neural network analysis was accomplished by TIBERIUS Data Mining version 7.0.7a (copyright © Tiberius Data Mining, Melbourne Australia) and back propagation ANN Tiberius XL for EXCEL (by Phil Brierley, Melbourne Australia).

2.3 Multiple Regression Analysis

Multiple regression analysis was carried out by utilizing GraphPad Instat version 3.06 (copyright © 1992-2003 GraphPad Software, GraphPad Software, Inc. 7825 Fay Avenue, Suite 230, La Jolla, CA 92037 USA) and Smith's Statistical Package version 2.80 (copyright © 1995-2005 by Gary Smith, Pomona College, Claremont, California 91711). ANOSIM (analysis of similarity) and cluster analysis was performed by PAST version 2.15 (copyright Hammer and Harper 1999-2012, University of Oslo, Sars gate1, 0562 Oslo, Norway).

2.4 Statistical Analysis

Additional statistical analysis was accomplished for numerical outliers by Grubbs' test (or ESD, extreme studentized deviate) using GraphPad calculators of stats, last seen July 2014 at http://www.graphpad.com/quickcalcs/. Determination of F and T tests, Mann-Whitney, Kolmogorov-Smirnov, Kruskal-Wallis, Wilcoxon, and paired tests was accomplished using PAST version 2.15. Summary statistics and correlation was performed by EXCEL (Microsoft Office Professional Plus 2013) and by Smith's Statistical Package. Coefficient of determination, R^2, is calculated using GraphPad Instat.

3. RESULTS AND DISCUSSION

3.1 Skin Absorption, Topical Drugs, Dermal Permeability Coefficient

The major routes of compound diffusion passage through dermal layers is described in (Fig. 1). Penetration by intercellular and transcellular process is far more predominant than by transappendageal [1]. Transappendageal is penetration by way of a skin appendage (i.e. hair follicles, sebaceous glands, and/or sweat glands); however, these appear on 1% of the human skin surface and their capacity as transport channels for the passage of compounds from the external environment to the capillary bed is most often negligible [1].

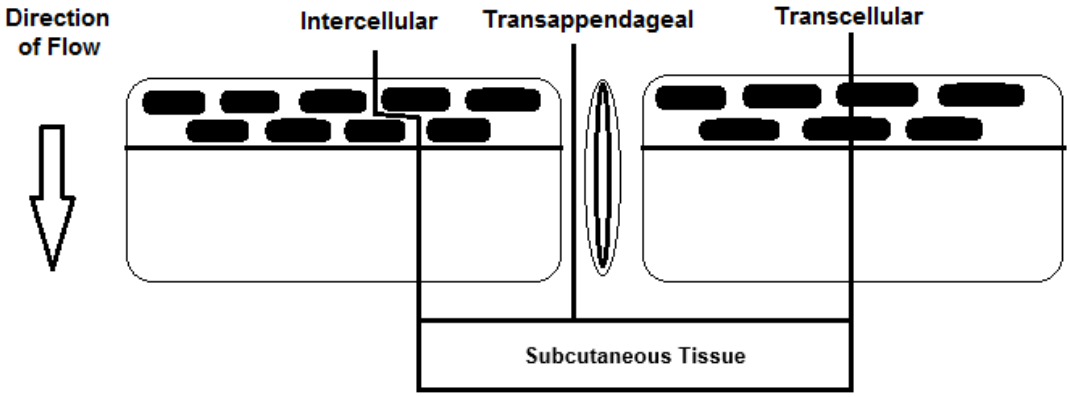

Fig. 1. Major routes of diffusion through the skin. Because appendages occupy less than 1% of the skin surface for humans, their role as transport channels for the passage of substances from the external environment to the capillary bed is generally negligible

Molecular descriptors are now playing a key role in scientific research and much work has been accomplished to establish quantitative relationships between structures and properties, biological activities and other experimental properties [18]. For this study the molecular constitutional descriptors [18] of molecular weight, number of oxygen & nitrogen atoms (these are hydrogen bond acceptors or nON), number of rotational bonds (nrotb), polar surface area (PSA), number of hydrogen bond donors (these are $-NH_n$, -OH or nNHOH), etc. are determined for a very broad category of medicinal compounds (see Table 1). In addition, molecular properties of Log $K_{o/w}$ and topological polar surface area are included in the characterization of this population of medicaments for purpose of demonstrating the efficacy of Kp computation by artificial neural network analysis and multiple regression.

To study the efficacy of predicting Kp values utilizing other descriptors with or without Log $K_{o/w}$ and molecular weight (MW), a broad range of drug categories is selected (see Table 1). Various categories included in (Table 1) are: 1) antimicrobial (drugs 1,2,3,4,5); 2) anticholinergic (drug 6); 3) antiemetic (drug 7); 4) psychostimulant (drug 8); 5) anesthetic (drug 9); 6) steroid hormone (drug 10); 7) stimulant (drug 11); 8) alkaloid drug (drug 12); 9) NO enhancer (drug 13); 10) hormone (drug 14); 11) adrenergic agonist (drug 15); 12) opoid agonist (drug 16); 13) corticosteroid (drugs 17, 18, 19); 14) analgesic (drug 20). Altogether (Table 1) is a highly diversified composite of drug categories, this in addition to their 12 descriptors of pharmacological significance.

Interestingly, this population of 20 diverse clinical drugs have properties that cause each drug to be in very high positive correlation to all others (Pearson r > 0.8000). That is, the 12 descriptors for each drug is very highly correlated to all others. In addition, the set of 12 descriptors for each drug (not by individual descriptors), the Kruskal-Wallis test indicates the 20 drugs have similar medians (P =1.0). Descriptor properties having a positive skewness (e.g. the mean and the median are both greater than the mode) are Kp, Log Kp, PSA, MW, nON, nOHNH, nrotb, molecular volume (MV), Rule of 5, and diffusion coefficient. Only Log $K_{o/w}$ has a negative skewness (the mean and the median are both less than the mode).

ANOSIM (analysis of similarity) outcome of R = 1.005 indicated that significant differences exist among the drug samples of (Table 1) (P = 1.0) based on numerical values of these properties. Cluster analysis shows additional information of this group of diverse drugs (see Fig. 2). Using pair-group average (clusters are joined by average distance between all membranes in group) and Euclidean distance (the ordinary distance between two points measured with a ruler).

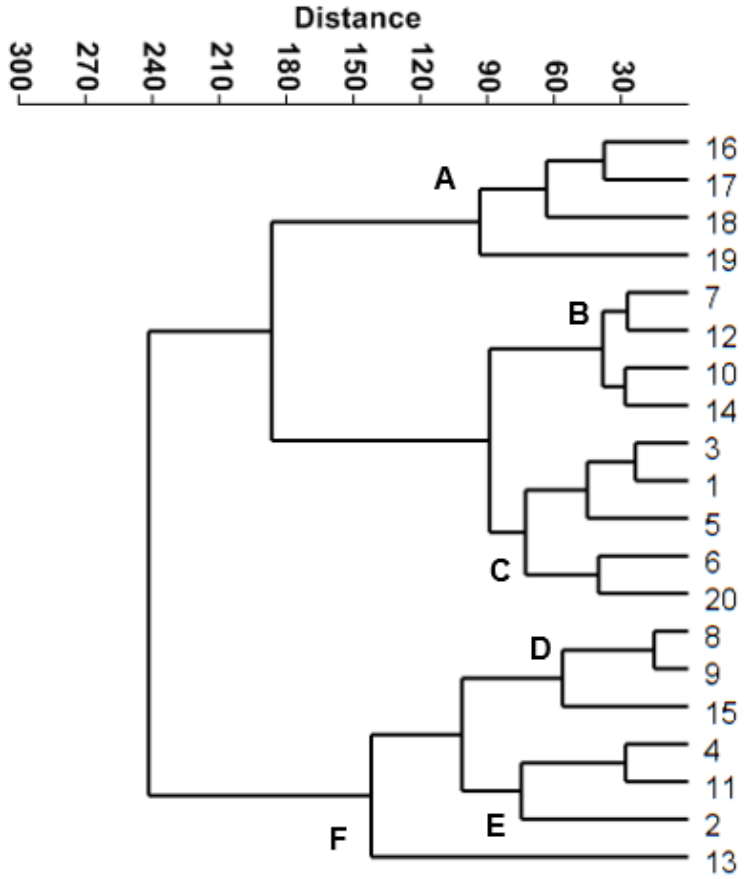

Fig. 2. Cluster analysis according to highest level of similarity. Drugs 16, 17, 18, 19 are most similar (joined at node A); drugs 7, 12, 10, 14 are most similar (joined at node B); drugs 3, 1, 5, 6, 20 are most similar (joined at node C); drugs 8, 9, 15 are most similar (joined at node D); and drugs 4, 11, 2 most similar (joined at node E); and 13 appears distinct (joined at node F)

The raw number of six clusters of 20 drugs total in the horizontal dendrogram (see Fig. 2) suggests significant differences within the multivariate table of molecular properties presented in (Table 1).

3.2 Prediction of Dermal Permeability Constant by Artificial Neural Network (ANN)

Artificial neural network modeling applying artificial intelligence are referred to as artificial neural networks (ANN) that for this study consists of three layers (see Fig. 3). Patterns (data) are presented to the network by means of the input layer, followed by communication to one or more hidden layers [19]. For this study the three layers can be represented as the input layer, hidden layers, then output layer where outcomes are presented. The ANN system applied in this study (Tiberius) is a feed forward multilayer perceptrion trained with a back propagation algorithm. This system is a supervised learning approach, where, in supervised learning a set of examples is provided and the goal is to find a function in the allowed class of functions that matches these examples. The result is a function to infer the mapping that is implied by the data [20]. Numerical analysis by ANN is a type of data mining technique and has been utilized for pharmacokinetic and pharmacodynamics modeling and analysis [21-24], as well as quantitative structure activity relationships [25]. ANN does not require rigid experimental designs and can complete mapping with historical or incomplete data by gathering information through experience from patterns in the data [25,26]

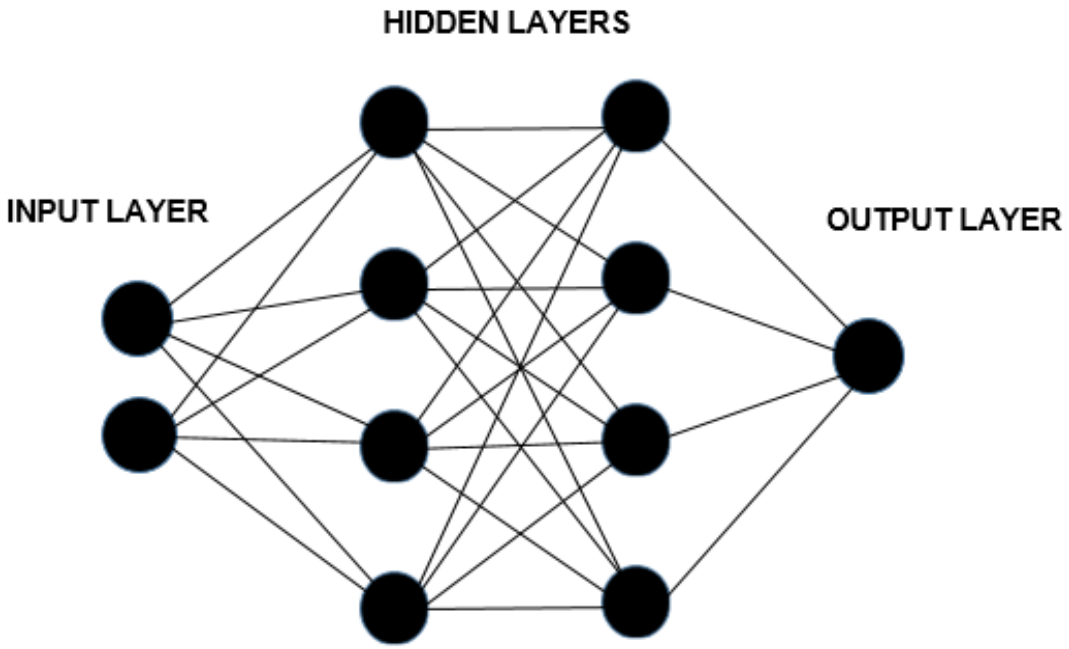

Fig. 3. Artificial neural networks are commonly organized in layers that are made up of interconnected nodes which contain an activation function. Patterns (data) are presented to the network by means of the input layer, followed by communication to one or more hidden layers. Within the hidden layers, the actual processing is done by utilizing a system of weighted connections. The hidden layers then connect to an output layer where the answer is output

Table 1. Properties of drugs

Drug	Kp (cm/hour)	Log Kp	Log KO/W	Polar surface area (A^2)	Natoms	MW	nON	nOHNH	Rotatable bonds	Molecular volume (angstroms^3)	Rule of 5	Diffusion coefficient (cm^2/second)
1 *Ofloxacin* (antimicrobial)	0.00000629	-5.201	-0.262	75.01	26	361.37	7	1	2	311.146	0	5.495E-15
2 Mafenide (antimicrobial)	0.0000564	-4.287	-0.184	86.19	12	186.24	4	4	2	154.85	0	4.418E-13
3 Gatifloxacin (antimicrobial)	0.0000221	-4.656	-0.036	83.8	27	375.4	7	2	4	327.59	0	6.783E-14
4 Chloroxylenol (antimicrobial)	0.0443	-1.354	2.889	20.228	10	156.61	1	1	0	138.72	0	2.726E-07
5 Ampicillin (antimicrobial)	0.000128	-3.893	-0.873	112.73	24	349.41	7	4	4	298.87	0	2.276E-12
6 Oxybutynin (anticholinergic)	0.00017	-3.77	4.862	49.771	26	357.49	4	1	8	361.79	0	4.014E-12
7 Granisetron (antiemetic)	0.0032	-2.494	2.428	50.162	23	312.42	5	1	2	299.35	0	1.422E-09
8 Methylphenidate (psychostimulant)	0.00648	-2.188	2.283	38.332	17	233.31	3	1	4	231	0	5.865E-09
9 Lidocaine (anesthetic)	0.00383	-2.417	2.13	32.336	17	234.34	3	1	5	244.86	0	2.037E-09
10 Testosterone (steroid hormone)	0.00755	-0.295	3.245	37.3	21	288.43	2	1	0	291.54	0	7.917E-09
11 *Nicotine* (stimulant)	0.00132	-2.88	1.09	16.13	12	162.24	2	0	1	165.62	0	2.420E-10
12 Scopolamiine (alkaloid drug)	0.000133	-3.876	1.046	62.3	22	303.36	5	1	5	277.21	0	2.457E-12
13 *Nitroglycerin* (no enhancer)	0.00111	-2.954	2.188	165.17	15	227.09	12	0	8	160.02	1	1.711E-10
14 Estradiol (hormone)	0.0292	-1.535	3.43	40.46	20	272.39	2	2	0	268.74	0	1.184E-07
15 *Clonidine* (adrenergic agonist)	0.00155	-2.81	2.612	36.42	14	230.1	3	2	2	181.94	0	3.337E-10
16 *Buprenorphine* (opioid agonist)	0.00919	-2.037	4.872	62.162	34	467.65	5	2	5	449.63	0	1.173E-08
17 Clobetasol propionate (corticosteroid)	0.000825	-3.084	3.678	80.675	32	467	5	1	5	416.9	0	9.453E-11
18 Betamethasone dipropionate (corticosteroid)	0.00123	-2.91	4.185	106.98	36	504.6	7	1	8	464.7	1	2.101E-10
19 *Halobetasol* (corticosteroid)	0.000162	-3.791	2.628	74.598	29	428.9	4	2	2	368.55	0	3.645E-12
20 Fentanil (analgesic)	0.0127	-1.896	3.7393	23.547	25	336.48	3	0	6	340.163	0	2.240E-08

ANN is applied in this study to determine the efficacy of obtaining Kp numerical values consistent with the established and recognized approach utilizing equation (1) [14-17]. Therefore, ANN will be applied to analyze the following models of descriptors: (A) Determination of Log Kp using Log $K_{o/w}$ and MW, only; (B) Determination of Log Kp by utilizing Log $K_{o/w}$, MW, PSA, natoms, nON, nrotb, nOHNH, and molecular volume descriptors; (C) Determination of Kp by utilizing PSA, natoms, nON, nrotb, nOHNH, and molecular volume descriptors; and (D) Determination of Kp utilizing Log $K_{o/w}$, MW, PSA, natoms, nON, nrotb, nOHNH, and molecular volume descriptors.

Deming regression is an important statistical method that should be used when comparing a known method and a test method [27-29]. Deming regression method, allows the comparison of two measurement methodologies (techniques for measuring Kp or Log Kp) which supposes that measurement error is present in both the X (known) and Y (test method).

A Deming regression line with a slope of one and an intercept of zero would indicate that the two methodologies give similar results throughout the entire range of measurements [27-29].

For rapid and efficient evaluation and visualization of outcome by ANN application for predicting Log Kp and Kp from multiple descriptor properties presented in (Table 1), the Deming regression plots are place adjacent to each other for all models (A), (B), (C), and (D) (see description above), shown in (Fig. 4). For Deming regression the identity line, where y=x having slope of 1.000, is not shown. The Pearson r correlation for the linear regression lines within the four plots are as follows (see Fig. 4): (A) r = 0.8147; (B) r = 0.8168; (C) r = 0.5510; and (D) r = 0.6972. Clearly, by values of r, for Deming regression models (A) and (B) in comparing of the actual Log Kp values (x-axis) to ANN predicted values (y-axis), have a very strong positive correlation to the ANN predicted values for Log Kp. Remembering that model (A) is the standard Log Kp relationship equation (1) using only descriptors Log $K_{o/w}$ and MW, it appears that ANN can predict Log Kp at a similar accuracy in model (B) with additional descriptors PSA, natoms, nON, nrotb, nOHNH, and molecular volume, along with Log $K_{o/w}$ and MW. In model (B) the investigator can incorporate additional properties that reflect important pharmaceutical characteristics. For example, the number of oxygen & nitrogen atoms (or nON) as well as the number of amine and hydroxyl groups (or nNHOH) are indicators of hydrogen bond acceptor and donor features of a lead drug which are used to evaluate drug-likeness. The ability to apply more property descriptors than only MW and Log $K_{o/w}$ should allow more extensive evaluation of a new lead compound as well as incorporate increased level of compound property descriptors that represent the molecular scaffold. The coefficient of determination (R^2) for model (A) and (B) indicates that 66.37% and 66.72% of the variance is represented by the model, respectively.

For model (A) statistical analysis of actual Log Kp to predicted Log Kp showed no outliers by Grubb's test and the following: 1) By t-test and Wilcoxon test the actual and predicted Log Kp values have equal means and median (P = .55 and P = .31, respectively); 2) By Mann-Whitney and Kolmogorov-Smirnov test the actual and predicted Log Kp values have equal medians and equal distribution (P = .54 and P = .27, respectively).

For model (B) statistical analysis of actual Log Kp to predicted Log Kp showed no outliers by Grubb's test and the following: 1) By t-test and Wilcoxon test the actual and predicted Log Kp values have equal means and median (P = .98 and P = .74, respectively); 2) By Kruskal-Wallis and Kolmogorov-Smirnov test the actual and predicted Log Kp values have equal medians and equal distribution (P = .91 and P = .97, respectively).

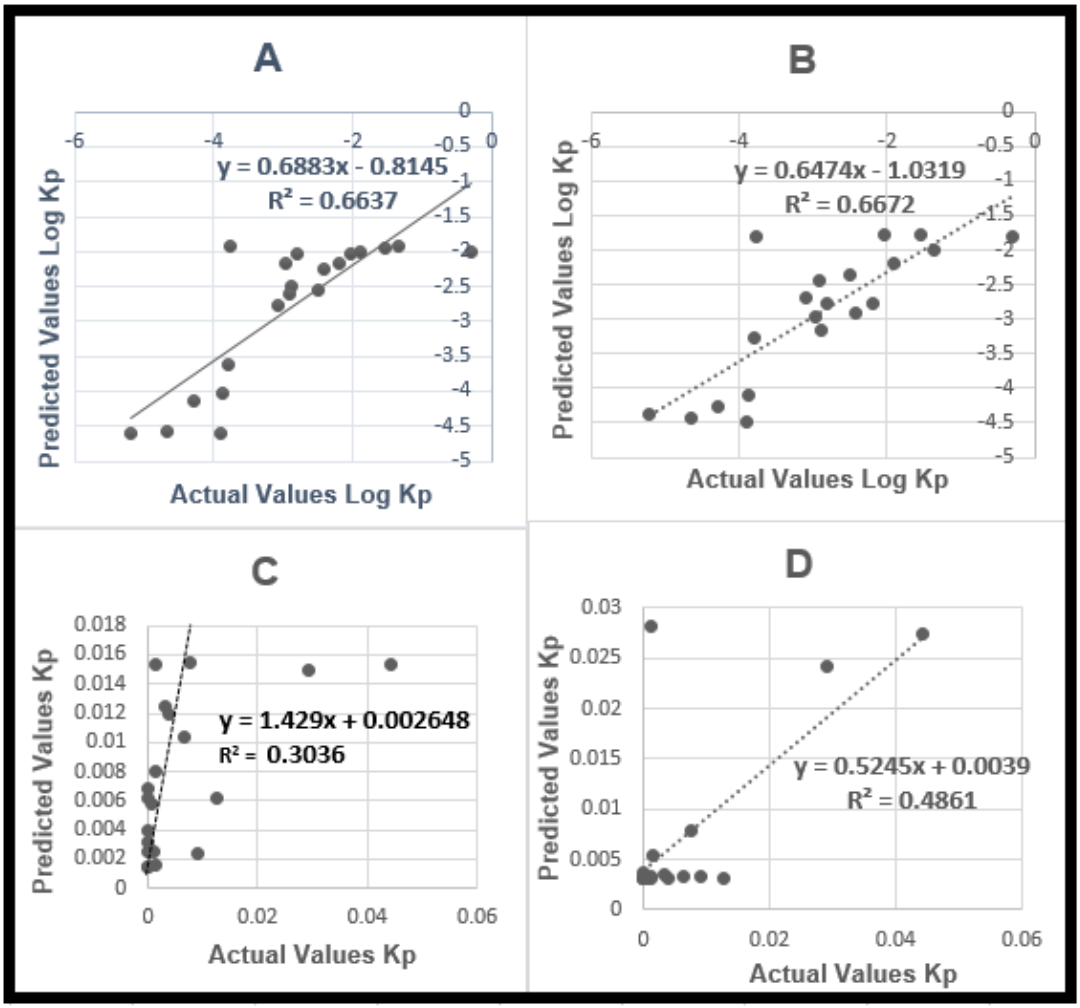

Fig. 4. Deming regression analysis and comparison of prediction of Log Kp and Kp by artificial neural network. Plot A is comparison of Log Kp values of actual values (x-axis) to predicted values (y-axis) utilizing Log Kow and MW only. Plot B is comparison of Log Kp values of actual values (x-axis) to predicted values (y-axis) utilizing Log Ko/w, MW, PSA, natoms, nON, nrotb, nOHNH, and molecular volume (MV) descriptors. Plot C is comparison of raw numerical Kp values predicted by ANN showing actual values (x-axis) to predicted values (y-axis) utilizing PSA, natoms, nON, nrotb, nOHNH, and molecular volume (MV) descriptors. Plot D is comparison of raw numerical Kp values predicted by ANN, showing actual values (x-axis) to predicted values (y-axis) utilizing Log Ko/w, MW, PSA, natoms, nON, nrotb, nOHNH, and molecular volume descriptors

Clearly, for Deming regression models (C) and (D) seeing the correlation r values for comparing of the actual Kp values (x-axis) to the ANN predicted values (y-axis), have a strong positive relationship to the ANN predicted values for Kp. Remembering that model (C) is the prediction of Kp using PSA, natoms, nON, nrotb, nOHNH, and molecular volume descriptors. Model (D) is the comparison of actual Kp values (x-axis) to ANN predicted Kp

values (y-axis) utilizing Log Ko/w, MW, PSA, natoms, nON, nrotb, nOHNH, and molecular volume descriptors. It appears that ANN has substantial difficulty of predicting the raw numerical values of Kp for both models. The slope value of 1.000 for the identity line (where y = x) compared to slope values of 1.429 (model (C)) and 0.5245 (model (D)), shows substantial deviation from the highest association of y = x. The R^2 for these models show 30.36% and 48.61% representation of the variance for (C) and (D), respectively. Clearly, ANN is not effective for predicting raw Kp numerical values even from increased number of property descriptors.

For model (C) statistical analysis of actual Kp to predicted Kp by by ANN showed one outlier by Grubb's test (for drug 4) and the following: 1) By t-test and Mann-Whitney test the actual and predicted Kp values have equal means but not equal medians (P = .56 and P = .01, respectively); 2) By Kruskal-Wallis and Kolmogorov-Smirnov test the actual and predicted Kp values do not have equal medians and do not have equal distribution (P = .01 and P = .003, respectively).

For model (D) statistical analysis of actual Kp to predicted Kp by by ANN showed one outlier by Grubb's test (for drug 4) and the following: 1) By t-test and Mann-Whitney test the actual and predicted Kp values have equal means but not equal medians (P = .61 and P = .03, respectively); 2) By Kruskal-Wallis and Kolmogorov-Smirnov test the actual and predicted Kp values do not have equal medians and do not have equal distribution (P = .01 and P = .003, respectively).

Much of the statistical analysis is suggested in (Table 2) showing summary statistics of Log Kp and Kp, actual and predicted by ANN. Note for models (A) and (B) the mean values are extremely close, however, for models (C) and (D) the mean values are distant from each other. Also, the medians for values of Log Kp and Kp for models (A) and (B) are extremely close, however, this is not the case in models (C) and (D).

In short, ANN is able to efficiently predict Log Kp in model (A) utilizing only the Log Ko/w and MW descriptors (same as equation (1)), but also in the model (B) utilizing Log Ko/w, MW, PSA, natoms, nON, nrotb, nOHNH, and molecular volume descriptors. However, ANN failed to adequately predict Kp numerical values in model (C) using descriptors PSA, natoms, nON, nrotb, nOHNH, and molecular volume; and model (D) utilizing Log Ko/w, MW, PSA, natoms, nON, nrotb, nOHNH, and molecular volume descriptors.

3.3 Prediction of Dermal Permeability Constant Utilizing Multiple Regression

There are two general applications for multiple regression (MR): prediction and explanation [27,28,30]. Using MR for explanatory purposes is exploring relationships among multiple variables of a model to elucidate a new understanding of a population. When one uses MR for prediction it is using a model to create a regression equation that would optimally predict a particular event within the population [30]. Regression equations obtained for cases (A), (B), (C), and (D) (see Fig. 4) will now be presented.

For prediction of Log Kp for 20 drug types (see Table 1) utilizing only Log Ko/w and MW for case (A), shown in (Fig. 4). The outcome produced the following regression function:

$$\text{Log Kp} = 0.01597 + 0.003456\,(\text{Log Ko/w}) - 0.00005768\,(\text{MW}) \qquad (2)$$

Table 2. Comparison of actual values of log Kp and Kp to predicted outcome using tiberius ANN

	A		B		C		D	
	Tiberius ANN prediction of Log K_p by Log $K_{O/W}$ AND MW		Tiberius ANN prediction of Log K_p by all properties with Log $K_{O/W}$ and MW		Tiberius ANN prediction of K_p by all properties (NO Log $k_{O/W}$, NO MW)		Tiberius ANN prediction of K_p by all properties with Log $k_{O/W}$ and MW	
	Actual Log Kp	Predicted Log Kp	Actual Log Kp	Predicted Log Kp	Actual Kp	Predicted Kp	Actual Kp	Predicted Kp
Mean	-2.916	-2.821	-2.916	-2.919	0.006158	0.007407	0.0074077	0.007109
Standard error	0.2691	0.2274	0.2691	0.2133	0.002530	0.001168	0.001168	0.001904
Median	-2.895	-2.372	-2.895	-2.775	0.001275	0.006215	0.006215	0.003221
Standard deviation	1.203	1.016	1.203	0.9540	0.011318	0.005224	0.005224	0.008515
Sample variance	1.448	1.034	1.448	0.9101	0.0001281	2.729E-05	2.729E-05	7.250E-05
Kurtosis	-0.0401	-0.8858	-0.0401	-1.059	7.0820	-1.338	-1.3384	2.771
Skewness	0.1314	-0.9039	0.1314	-0.4989	2.6548	0.45695	0.45695	2.0841
Range	4.906	2.693	4.906	2.704	0.04429	0.01410	0.014106	0.02497
Minimum	-5.201	-4.605	-5.201	-4.488	0.00000629	0.001418	0.001418	0.003205
Maximum	-0.295	-1.912	-0.295	-1.784	0.0443	0.01552	0.01552	0.02817
Sum	-58.38	-56.43	-58.32	-58.39	0.12316	0.14815	0.1482	0.14217
Count	20	20	20	20	20	20	20	20

The computation package determining this equation showed that the constant as well as both variable were significant and contribute significantly to the model. Both the Log Ko/w and MW data are normally distributed. An R^2 of 0.3057 is reported, indicating that 30.57% of variation in Log Kp results can be explained by the regression. The coefficient of determination or R^2, indicates how well data fit a statistical model and is a statistic used in statistical models whose main purpose is either the prediction of future outcomes or the testing of hypotheses [27,28]. The R^2 value provides a measure of how well observed outcomes are replicated by the model [27,28]. This is a low value for R^2, to be compared to the following approaches utilizing additional/different descriptors.

For prediction of Log Kp for 20 drug types utilizing descriptors for the case (B), shown in (Fig. 4) (descriptors: PSA, natoms, nON, nOHNH, nrotb, Log Ko/w, MW, and molecular volume (MV)). The outcome produced the following regression function:

$$\text{Log Kp} = 0.02872 + 0.0001658(\text{PSA}) + 0.01557(\text{natoms}) - 0.006442(\text{nON}) + 0.002614(\text{nOHNH}) + 0.0001645\ (\text{nrotb}) - 0.0008177(\text{MV}) + 0.006574\ (\text{Log Ko/w}) - 0.0004227(\text{MW}) \quad (3)$$

The computation package determining this equation showed that only the constant as well as Log Ko/w were significant and contribute significantly to the model. Numbers for all descriptors, except Log Kp and nOHNH were normally distributed. An R^2 of 0.5875 is reported, indicating that 58.75% of variation in Log Kp results can be explained by the regression. This is a much higher value R^2 than found for case (A) above.

For prediction of Kp for 20 drug types utilizing descriptors for the case (C), shown in (Fig. 4) (descriptors: PSA, natoms, nON, nOHNH, nrotb, and molecular volume (MV)). The outcome produced the following regression function:

$$\text{Kp} = 0.02474 + 0.0001562(\text{PSA}) + 0.0008644\ (\text{natoms}) - 0.003955\ (\text{nON}) - 0.002962\ (\text{nOHNH}) - 0.0009285\ (\text{nrotb}) - 0.00007585\ (\text{MV}) \quad (4)$$

The computation package determining this equation showed that none of the descriptors or constant contribute significantly to the model. Numbers for all descriptors, except Kp and nOHNH were normally distributed. An R^2 of 0.3346 is reported, indicating that 33.46% of variation in Kp results can be explained by the regression. This is a low value for R^2, to be compared to the other approaches utilizing additional/different descriptors.

For prediction of Kp for 20 drug types utilizing descriptors for the case (D), shown in (Fig. 4) (descriptors: PSA, natoms, nON, nOHNH, nrotb, MV, Log Ko/w, and MW). The outcome produced the following regression function:

$$\text{Kp} = 0.02872 + 0.0001658(\text{PSA}) + 0.01557(\text{natoms}) - 0.006442(\text{nON}) + 0.002614(\text{nOHNH}) + 0.0001646(\text{nrotb}) - 0.0008177(\text{MV}) + 0.006574(\text{Log Ko/w}) - 0.0004227(\text{MW}) \quad (5)$$

The computation package determining this equation showed that the constant and Log Ko/w contribute significantly to the model. Numbers for all descriptors, except Kp and nOHNH were normally distributed. An R^2 of 0.5875 is reported, indicating that 58.75% of variation in Kp results can be explained by the regression. This is a much higher value of R^2 than found in case (C) above.

Clearly, the use of multiple regression produces effective results in cases of highest R^2. Essentially, multiple regression prediction for Log Kp is most effective when utilizing

descriptors for case (B), shown in (Fig. 4) (descriptors: PSA, natoms, nON, nOHNH, nrotb, Log Ko/w, MW, and MV). Than the R^2 becomes 0.5875 and represents 58.75% of variance found within the model. The model of case (A) (use of descriptors Log Ko/w and MW) failed to produce a substantial R^2.

Essentially, multiple regression prediction for Kp is most effective when utilizing descriptors for case (D), shown in (Fig. 4) (descriptors: PSA, natoms, nON, nOHNH, nrotb, MV, Log Ko/w, and MW). Than the R^2 becomes 0.5875 and representing 58.75% of variance found within the model. The model of case (C) (use of descriptors: PSA, natoms, nON, nOHNH, nrotb, and MV) failed to produce a substantial R^2.

Increasing the number of property descriptors to construct the regression model tends to improve the R^2 and increase the extent that the variance of the model is represented by the regression equation. The increased number of descriptors allows the investigator to increase the level of pharmacological important properties into a study and prediction of the permeability coefficient (Kp), however, the R^2 remains difficult to produce to values greater than 0.8000.

4. CONCLUSION

In summation, 20 drugs that cover a broad spectrum of 12 categories of drugs, had numerous properties determined and analyzed by cluster analysis to show a total of five main clusters of similarity. Applying a forward feeding back propagation artificial neural network analysis for determination of Log Kp and Kp showed mixed results.

Outcome of ANN analysis for Log Kp and Kp determination:

1. ANN analysis was able to effectively predict Log Kp in the case utilizing Log Kow and MW.
2. ANN analysis was able to effectively predict Log Kp in the case utilizing Log Kow, MW, PSA, natoms, nON, nrotb, nOHNH, and molecular volume (MV) descriptors.
3. ANN analysis failed to produce accurate numerical values of Kp utilizing descriptors PSA, natoms, nON, nrotb, nOHNH, and molecular volume (MV) descriptors.
4. ANN analysis failed to produce accurate numerical values of Kp utilizing descriptors Log Ko/w, MW, PSA, natoms, nON, nrotb, nOHNH, and molecular volume descriptors.

For comparison, a multiple regression analysis for Log Kp and Kp utilizing various number of descriptors produced mixed results, when considering R^2 and the amount of variance within the model that is represented by the regression equation.

Outcome for prediction of Log Kp and Kp determination by multiple regression analysis:

1. Multiple regression failed to produce a regression model for prediction of Log Kp utilizing Log Kow and MW descriptors ($R^2 = 0.3057$).
2. Multiple regression improved substantially in predicting Log Kp from an increased number of descriptors (PSA, natoms, nON, nOHNH, nrotb, Log Ko/w, MW, and molecular volume (MV)) ($R^2 = 0.5875$).
3. Multiple regression failed to produce a regression model for prediction of Kp utilizing PSA, natoms, nON, nOHNH, nrotb, and molecular volume (MV)) ($R^2 = 0.3346$).

4. Multiple regression improved substantially in predicting Kp utilizing additional descriptors PSA, natoms, nON, nOHNH, nrotb, MV, Log Ko/w, and MW) (R^2 = 0.5875).

The ability to determine Kp by using a variety of molecular properties will be useful for discerning the effects that control the efficiency of skin penetration by complex and sophisticated drug scaffolds. Design and development of novel topical drugs will be aided by applying the dermal permeability coefficient based on multiple pharmacological properties.

CONSENT

Not applicable.

ETHICAL APPROVAL

Not applicable.

COMPETING INTERESTS

Author has declared that no competing interests exist.

REFERENCES

1. United States Environmental Protection Agency. Dermal Exposure Assessment: Principles and Applications. 1992. EPA/600/8-91/011B.
2. Schultz SG. Basic Principles of Membrane Transport. Cambridge: Cambridge University Press; 1980.
3. Steen-Knudsen O. Biological Membranes. Theory of transport, potentials and electric impulses. Cambridge: Cambridge University Press; 2002.
4. Gennaro AR. Remington: The science and practice of pharmacy. 19th ed. Easton: Mack Publishing Company; 1995.
5. Bajaj S, Whiteman A, Brandner B. Parmacokinetics of transdermal drug delivery. Disclosures Cont Edu Anaseth Crit Care and Pain. 2011;11(2):39-43.
6. Paudel K, Milewski M, Swadley C, Brogden N, Ghosh P, Stinchcomb AL. Challenges and opportunities in dermal/transdermal delivery. Ther Deliv. 2010;1(1):109-131.
7. Wilhelm KP, Zhai H, Maibach HI. Dermatotoxicology. New York: CRC Press; 2007.
8. Kwatra S, Taneja G, Nasa N. Alternatives routes of drug administration-transdermal, pulmonary & parenteral. Indo Global Journal of Pharmaceutical Sciences. 2012;2(4):409-26.
9. Margetts L, Sawyer R. Transdermal drug delivery: principles and opioid therapy. Cont Edu Anaseth Crit Care and Pain. 2007;7(5):171-76.
10. Fasano WJ, McDougal JN. In vitro dermal absorption rate testing of certain chemicals of interest to the occupational safety and health administration: summary and evaluation of USEPA's mandated testing. Regul Toxicol Pharmacol. 2008;51(2):181-94.
11. Rutter N. Drug absorption through the skin: A mixed blessing. Archives of Disease in Childhood. 1987;62:220-21.
12. Montagna W, Van Scott EJ, Stoughton RB. Pharmacology and the skin. New York: Appleton Century Crofts; 1972.
13. Shaw JE, Urquhart J. Transdermal drug absorption: A nuisance becomes an opportunity. Br Med J. 1981;283:875-6.

14. Kupczewska-Dobecka M, Jakubowsk M, Czerczak S. Calculating the dermal flux of chemicals with OELs based on their molecular structure: An attempt to assign the skin notation. Environmental toxicology and pharmacology. 2010;30:95-102.
15. Sun Y, Lam LT, Moss G, Prapopoulou M, Adams R, Davey N, et al. Predicting drug absorption rates through human skin. IEEE International Joint Conference on Neural Networks (IJCNN); 2010. No.11593964.IEEE:1-5.
16. Roberts M. Dermal absorption and toxicity assessment. 2n ed. New York: CRC Press; 2007.
17. Potts RO, Guy RH. Predicting skin permeability. Pharm Res. 1992;9:663-69.
18. China Computational Biology Drug Design Group. Molecular Descriptors Guide. Accessed 10 July 2014.
Available: http://www.google.com/search?q=molecular+descriptors+guide&ie=utf-8&oe=utf-8&aq=t&rls=org.mozilla:en-US:official&client=firefox-a&channel=sb&gws_rd=ssl.
19. Livingstone DJ. Artifical neural networks methods and applications. New York: Humana Press; 2009.
20. Cheng F, Sutariya V. Applications of artificial neural network modeling in drug discovery. Clin Exp Pharmacol. 2012;2(3):e113.
21. Haidar SH, Johnson SB, Fossler MJ, Hussain AS. Modeling the pharmacokinetics and pharmacodynamics of a unique oral hypoglycemic agent using neural networks. Pharm Res. 2002;19:87-91.
22. Gobburu JV, Chen EP. Artificial neural networks as a novel approach to integrated pharmacokinetic-pharmacodynamic analysis. J Pharm Sci. 1996;85:505-510.
23. Chow HH, Tolle KM, Roe DJ, Elsberry V, Chen H. Application of neural networks to population pharmacokinetic data analysis. J Pharm Sci. 1997;86(7):840-5.
24. Ritschel WA, Akileswaran R, Hussain AS. Application of neural networks for the prediction of human pharmacokinetic parameters. Methods Find Exp Clin Pharmacol. 1995;17:629-643.
25. Sutariya V, Groshev A, Sadana P, Bhatia D, Pathak Y. Artificial neural network in drug delivery and pharmaceutical research. The Open Bioinformatics Journal. 2013;7(suppl1,M5):49-62.
26. Agatonovic-Kustrin S. Basic concepts of artificial neural network (ANN) modeling and its application in pharmaceutical research. J Pharm Biomed Anal. 2000;22(5):717-727.
27. Cohen J, Cohen P. Applied multiple regression/correlation analysis for the behavioral sciences. Hillsdale: Lawrence Erlbaum Associates; 1983.
28. Pedhazur EJ. Multiple regression in behavioral research. Orlando: Harcourt Brace; 1997.
29. Tabachnick BG, Fidell LS. Using multivariate statistics. New York: Harper Collins; 1996.
30. Anderson N H, Shanteau J. Weak inference with linear models. Psychological Bulletin. 1977;84:1155-1170.

Peer-review history:
The peer review history for this paper can be accessed here:
http://www.sciencedomain.org/review-history.php?iid=637&id=22&aid=6045