Linking existing *in vitro* dermal absorption data to physicochemical properties: contribution to the design of a weight-of-evidence approach for the safety evaluation of cosmetic ingredients with low dermal bioavailability

Gamze Ates[1*$], Fabian P. Steinmetz[2*], Tatyana Yordanova Doktorova[3], Judith C. Madden[2**], Vera Rogiers[1**]

[1]Department of In Vitro Toxicology and Dermato-cosmetology, Vrije Universiteit Brussel, Laarbeeklaan 103, B-1090 Brussels, Belgium.

[2] QSAR and Modelling Group, School of Pharmacy and Biomolecular Sciences Liverpool John Moores University, Byrom Street, Liverpool, L3 3AF, England

[3] Unit of Toxicology, Scientific Institute of Public Health (IPH), Juliette Wytsmanstraat 14, B-1050, Brussels, Belgium.

* G. Ates and F. Steinmetz equally contributing first author

** J. Madden and V. Rogiers equally contributing last author

[$]Corresponding author: Gamze.Ates@vub.ac.be

Abbreviations

A, (daily amount of product exposed to per kg body weight); CDK, (Chemistry Development Kit); DA, (dermal absorption); DE, (Directorate-General); KNIME, (Konstanz Information Miner); log P, (logarithm of the octanol:water partition coefficient); MOE, (Molecular Operating Environment); MoS, (margin of safety); MP, (melting point); MW, (molecular weight); NO(A)EL, (no observable (adverse) effect level); SCCS, (Scientific Committee on Consumer Safety); SED, (systemic exposure dose); SMILES, (simplified molecular-input line-entry specification); TPSA, (topological polar surface area); TTC, (threshold of toxicological concern); WHO, (World Health Organisation)

Keywords

In silico prediction, dermal absorption, modelling, bioavailability, risk assessment, cosmetic ingredient, physicochemical properties, safety evaluation

Abstract

To characterize the risk of cosmetic ingredients when threshold toxicity is assumed, often the "margin of safety" (MoS) is calculated. This uncertainty factor is based on the systemic no observable (adverse) effect level (NO(A)EL) which can be derived from *in vivo* repeated dose toxicity studies. As *in vivo* studies for the purpose of the cosmetic legislation are no longer allowed in Europe and a validated *in vitro* alternative is not yet available, it is no longer possible to derive a NO(A)EL value for a new cosmetic ingredient. Alternatively, cosmetic ingredients with a low dermal bioavailability might not need repeated dose data, as internal exposure will be minimal and systemic toxicity might not be an issue. This study shows the possibility of identifying compounds suspected to have a low dermal bioavailability based on their physicochemical properties (molecular weight, melting point, topological polar surface area and log P) and their *in vitro* dermal absorption data. Although performed on a limited number of compounds, the study suggests a strategic opportunity to support the safety assessor's reasoning to omit a MoS calculation and to focus more on local toxicity and mutagenicity/genotoxicity for ingredients for which limited systemic exposure is to be expected.

Introduction

According to the European Cosmetic Regulation (EC 1223/2009), every cosmetic product on the market has to be safe for human health. This safety is based on the safety of its composing ingredients, their chemical structure, toxicological profile and exposure pattern. To characterize the risk of a cosmetic ingredient when threshold toxicity is assumed, the calculation of a so-called "margin of safety" (MoS) is applied. This uncertainty factor is used to extrapolate from test animals to humans and takes into account the systemic no observable (adverse) effect level (NO(A)EL) and the systemic exposure dose (SED). The former is derived either from *in vivo* oral repeated dose toxicity studies or reproductive toxicity data. The SED is estimated by taking into account the concentration (C) of the ingredient in the product, the daily amount of product exposed to per kg body weight (A, derived from consumer studies) and the dermal absorption (DA) [$\mathrm{MoS} = \frac{\mathrm{NOAEL_{sys}}}{\mathrm{SED}}$; SED (mg/kg bw/day) = A(mg/kg bw/day) x C(%)/100 x DA(%)/100]. As proposed by the World Health Organisation (WHO) an ingredient with a MoS $\geq$ 100 is considered to be safe (SCCS/1564/15).

However, with the introduction of the animal testing and marketing bans in the European cosmetic legislation and due to the absence of validated *in vitro* replacement methods for repeated dose or reproductive toxicity studies, it is no longer possible to derive a NO(A)EL to calculate the MoS for newly developed cosmetic ingredients. The consequences of this legal implementation start to become visible as no new UV-filters, preservatives or other cosmetic active ingredients have emerged in the last 2 years. So far only substances for which *in vivo* repeated dose studies were carried out before March 2013 have been evaluated by the SCCS. But for some particular ingredients the safety assessment might not be jeopardised. Indeed, ingredients with a negligible dermal bioavailability do not necessarily need repeated dose data, as internal exposure would be minimal and systemic toxicity might not be a potential issue. Adding the assumption that the main route of exposure to a cosmetic product is dermal and the dermal bioavailability will be in most cases even lower than the oral bioavailability, it might be justifiable to base the safety assessment of such

compounds on local toxicity and mutagenicity/genotoxicity test results, this on the assumption that no bioaccumulation is expected. In this context it is important to define when an ingredient is considered to have a low bioavailability.

Bioavailability, defined as the fraction of the dose administered (orally, dermally or via another route) that reaches the systemic circulation unchanged, is a composite parameter dependent on both absorption from the site of administration and metabolism of the compound. Within the area of drug discovery much research has been carried out on predicting bioavailability, particularly with respect to oral administration. As absorption is a key component, simple rules have been established that can be used to indicate the likelihood of absorption from the gastro-intestinal tract; the most well-known of these being the Lipinski rules (Lipinski et al., 2001). Briefly, the Lipinski "rule of fives" states that a logarithm of the octanol:water partition coefficient (log P) >5; molecular weight (MW) > 500; number of hydrogen bond acceptors >10; and number of hydrogen bond donors > 5 are features associated with poor oral absorption. Veber et al (2002) showed that compounds with high topological polar surface area (TPSA) and a high number of rotatable bonds are also associated with poor oral absorption. Furthermore a relationship has been shown between oral absorption and melting point (MP): chemicals with a higher MP are less likely to be absorbed (Chu et al. 2009). The "General Solubility Equation" relates melting point to solubility and partition coefficient. The advantage of using MP is that it is more easily determined than oral absorption. Additionally, Newby et al. (2015) have recently published decision trees to characterise the roles of certain physicochemical properties on the prediction of oral absorption. Whilst these rules are broad, and many exceptions are known, they demonstrate the principle that these simple physicochemical descriptors may be useful in classifying compounds as to high or low (oral) absorption.

Similarly, models have been developed to predict the extent of dermal penetration based on simple physicochemical properties; the most notable example being the work of Potts and Guy (1992) who demonstrated a correlation between log P and MW with skin permeability. Refinements to the Potts

and Guy model have since been published and the inherent difficulty modelling skin permeability data has been acknowledged (Steinmetz et al 2015). One problem is that measurements of dermal uptake are associated with high experimental variability, due to differences in assay conditions (e.g. differences in test protocols, skin type, use of solvents/vehicles, etc) making the development of robust, reliable quantitative models challenging. Another complication in modelling absorption is bias within the datasets. As skin is an effective barrier, dermal absorption data tend to be highly skewed towards low dermal absorption. The converse is observed for oral absorption data, as most is derived from drug development where high oral absorption is desirable, consequently most publically available data are for high oral absorption compounds. Despite these challenges, it would clearly be beneficial if rules based on simple physicochemical descriptors could be used to identify, accurately, compounds with low dermal absorption and thus low dermal bioavailability.

In order to investigate this possibility, a retrospective analysis of available safety evaluation data of cosmetic ingredients could provide valuable information. Although the safety of cosmetic products and their ingredients in Europe has to be assured by the companies' responsible person, for ingredients with some concern for human health *i.e.* colorants, preservatives, UV-filters and hair dyes (*i.e.* Annex substances), industry has to submit a full toxicological dossier to Directorate-General (DG) Sante when a mandate has been issued by DG Grow (Directorate-General for Internal Market, Industry, Entrepreneurship and SMEs of the EU). Risk assessment is than carried out according to the SCCS' Notes of Guidance (SCCS/1564/15). The resulting risk assessments, known as "opinions", are publically available via:
http://ec.europa.eu/health/scientific_committees/consumer_safety/index_en.htm. These contain summaries of the studies on the different toxicological endpoints as well as the physicochemical characteristics and DA of the ingredient under investigation. Data collated from these opinions can provide a high quality dataset for analysis. To formulate rules to identify low bioavailability compounds we undertook an empirical analysis of the cosmetic ingredients, assessed between 2000 and 2014 by the SCCS and its predecessors, to investigate the link between DA measured *in vitro* and

their physicochemical properties. In this study we propose a pragmatic approach that might aid in assessing whether a new cosmetic ingredient is likely to have a low dermal bioavailability.

Method

When preparing the data of all compounds from the SCCS opinions for modelling, the following criteria were used:

(i) DA measurements obtained using rat skin were excluded because of the relatively high uptake when compared to human or porcine skin.
(ii) if more than one DA measurement per compound were available an arithmetic mean was calculated.
(iii) descriptors were obtained for the parent form of the compounds.

A simplified molecular-input line-entry specification (SMILES) string for each compound was entered into the Molecular Operating Environment (MOE) software (version 2011.10) and processed to derive the neutralised form for the organic component. Topological polar surface area (TPSA) and molecular weight (MW) were calculated using a Chemistry Development Kit (CDK) node (molecular properties) available via the Konstanz Information Miner (KNIME) platform (KNIME version 2.10). The octanol:water partition coefficient (log P) was calculated using KowWin® (v1.68 available within EPI Suite 4.1, US EPA). Melting points (MPs) were extracted from the SCCS reports where possible. As the MPs of salts differ significantly from the MP of the parent compound, data was only included if available for the parent and not a salt form. This led to the creation of the data set (n = 70) used for further analysis.

For this data set a series of rules was defined in order to classify compounds as having a high or low DA. For this purpose a preliminary investigation was done using a Decision Tree Builder (KNIME version 2.10), employing log P, MW, TPSA and MP to determine which descriptors performed better in classifying compounds as high or low DA (data not shown here). Although the results from the decision tree alone were not conclusive, they provided guidance on the key descriptors, and

appropriate cut-off values that could be used to distinguish between high and low DA compounds. Based on these preliminary investigations, compounds were initially split into classes of low DA (<1.3%) or high DA (≥1.3%). The value of 1.3% was empirically derived to enable clear distinction between the classes. Because of the skew in % DA values (*i.e.* the majority of compounds have a low DA) the $\log_{10}$ of the DA values was used for ease of visualisation. Compounds showing greater DA have greater potential to induce systemic toxicity and therefore chemical features associated with higher percentage of DA are referred to here as physicochemical "alerts". The performance of the rules for the data set was investigated by calculating the sensitivity (correct classification of compounds with high DA) and specificity (correct classification of compounds with low DA). Rigid and flexible implementation of the derived rules was applied to optimise sensitivity and specificity of the results.

Results

The following physicochemical alerts, derived from preliminary investigation, were applied: (i) MW < 180 Da, (ii) log P ≥ 0.3, (iii) MP < 100°C, (iv)TPSA < 40 Å^2.

Compounds with MW < 180 Da and/or log P ≥ 0.3 and/or MP < 100°C and/or TPSA <40 Å^2 are more likely to be dermally absorbed. TPSA correlates with hydrogen bonding ability and preliminary investigations indicated that TPSA performed better than counts of hydrogen bond donors/acceptors in modelling the data here. The results are illustrated in figure 1, which shows that, in general, as the number of alerts increases the DA increases.

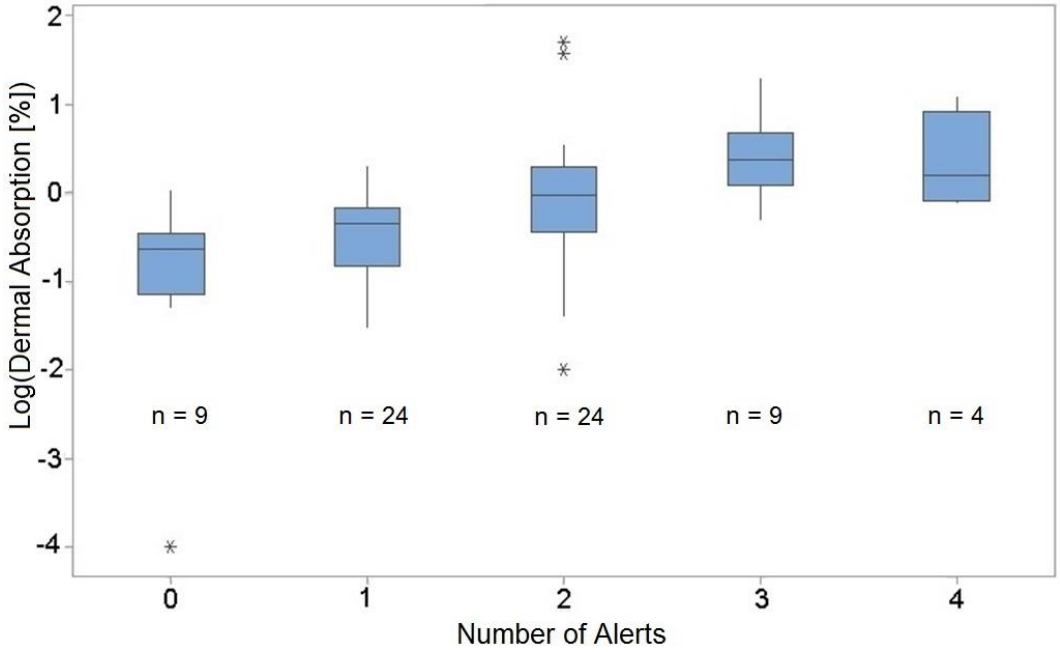

Figure 1: Boxplot of $\log_{10}$ % DA versus number of alerts for the data set (n=70). The alerts applied here are: MW <180 Da, log P ≥ 0.3, MP <100°C and TPSA < 40 Å^2. (MW= molecular weight, log P = octanol:water partition coefficient, MP = melting point, TPSA = topological polar surface area, * = outlier), horizontal bars indicate the median.

These rules can be interpreted as follows:

If any of the following criteria applies: (i) MW < 180 Da, (ii) log P ≥ 0.3, (iii) MP < 100°C and/or (iv) TPSA < 40 Å^2, then the compound is predicted as highly absorbed. If none of the criteria applies, the compound is predicted as poorly absorbed. Table 1 summarises the results of applying this rule set to the data set (n=70).

Table 1: Performance of the rule set on the data set (n=70). The number of compounds in each category is given, with the percentage of the total data set between brackets.

n = 70	Predicted High Absorption	Predicted Low Absorption	Total
High Absorption (≥1.3%)	23 (32.9%)	0 (0%)	23 (32.9%)
Low Absorption (<1.3%)	38 (54.3%)	9 (12.9%)	47 (67.1%)
Total	61 (87.1%)	9 (12.9%)	70 (100%)

The rule set shows a high sensitivity of 100% for the data set (*i.e.* for all 23 compounds in the high DA class all 23 were correctly predicted as being highly absorbed). The specificity of the rules is low as 38 out of 47 low DA compounds were incorrectly classified as highly absorbed rendering a specificity of 19.1%.

The results show that for the compounds studied here, when the rules predict a compound as having a low DA then the compound is likely to be poorly absorbed (no false negatives were identified using these rules). However, when the rules predict a compound as having a high absorption, then the compound may in fact have either a high or a low DA.

Flexible analysis of the data set

The same rule set was again applied to the same data set, however in this case additional flexibility was introduced. When a compound triggered none or only 1 of the alerts then it would still be predicted as low DA. Only compounds triggering two or more alerts would be assigned to the high absorption class. Table 2 shows the results for the data using the rule set with this more flexible interpretation.

Table 2: Performance of the rule set on the data set (n=70) with flexible interpretation (*i.e.* compounds must trigger two alerts to be placed in the high absorption class). The number of compounds in each category is given, with the percentage of the total data set between brackets.

n = 70	Predicted High Absorption	Predicted Low Absorption	total
High Absorption (≥1.3%)	19 (27.1%)	4 (5.7%)	23 (32.9%)
Low Absorption (<1.3%)	18 (25.7%)	29 (41.4%)	47 (67.1%)
total	37 (52.9%)	33 (47.1%)	70 (100%)

Table 2 shows that application of the rule set with flexible interpretation (*i.e.* two or more alerts need to be triggered to classify the compound as high DA) leads to an increased specificity (61.7%), but to a decreased sensitivity of 82.6% (*i.e.* 4 high DA compounds are now predicted as low DA). The increase in specificity may be out-weighed by the loss of sensitivity, as greater "cost" is associated with a false negative (*i.e.* predicting a high absorption compound as low DA).

However, it was noted that the 4 compounds that had been incorrectly classified into the low DA class all had a DA of < 2%. For this reason the analysis of the data set was repeated but in this case new boundaries were set for the two classes *i.e.* compounds for which DA was ≥ 2% were classified as high DA compounds, whereas those with DA <2% were taken as low DA compounds.

Results with new boundary criteria

The same rule set was applied to the data set, but with the cut-off value between high and low DA being set at 2%. Tables 3 and 4 show the results of applying the new cut-off value. In table 3 a compound is considered to belong to the class of high DA compounds if one or more alerts are triggered. In Table 4 the rule set is applied more flexibly and a compound is considered to belong to the high DA class only if two or more alerts are triggered.

Table 3: Performance of the rule set on the data set (n=70); triggering one or more alerts indicates a high DA compound. The number of compounds in each category is given, with the percentage of the total data set between brackets.

n=70	Predicted High Absorption	Predicted Low Absorption	Total
High Absorption (≥2%)	13 (18.6%)	0 (0%)	13 (18.6%)
Low Absorption (<2%)	48 (68.6%)	9 (12.9%)	57 (81.4%)
Total	61 (87.1%)	9 (12.9%)	70 (100%)

Table 4: Performance of the more flexible rule set on the data set (n=70); triggering two or more alerts indicates a high DA compound. The number of compounds in each category is given, with the percentage of the total data set between brackets.

n=70	Predicted High Absorption	Predicted Low Absorption	Total
High Absorption (≥2%)	13 (18.6%)	0 (0%)	13 (18.6%)
Low Absorption (<2%)	24 (34.3%)	33 (47.1%)	57 (81.4%)
Total	37 (52.9%)	33 (47.1%)	70 (100%)

The results given in tables 3 and 4 show that, when a cut-off value for DA of 2% is used, the sensitivity of the prediction is 100% in both cases. Indeed, compounds of high DA are always classified as highly absorbed; there are no false negatives. Allowing for a more flexible interpretation of the rule set, *i.e.* that 2 or more alerts need to be triggered in order for the compound to be predicted as having a high DA, increases the specificity – fewer true low DA compounds are predicted as having a high DA.

In summary, using the rule set with a cut-off value of 2% will lead to high DA compounds always being predicted as high (for this data set). However compounds with true low DA may be predicted as either high or low. More of the true low DA compounds are correctly classified when the more flexible rules are applied (specificity has increased from 15.8% to 57.9%).

Discussion

For the vast majority of cosmetic products the dermal route is the main route of human exposure. Therefore DA is a crucial factor in assessing the systemic toxicity of cosmetic ingredients. Another determining factor is the NO(A)EL. This value is in most cases derived from long-term *in vivo* repeated dose toxicity studies. Both factors are incorporated in the calculation of the MoS, an uncertainty factor to extrapolate from animals to humans (SCCS/1564/15). However, if evidence suggests that the compound under investigation has a low dermal bioavailability and thus systemic exposure is minimal, one might consider omitting the assessment of systemic toxicity. In the light of the animal testing and marketing bans of the European Cosmetic Regulation this would imply that data derived from an *in vivo* repeated dose toxicity study, for which no *in vitro* alternative yet exists, might not be needed for compounds with a negligible dermal bioavailability. In this context, it is important to define when a compound is considered to have a negligible dermal bioavailability. As described in the introduction, it is generally acknowledged that certain physicochemical properties such as MW, MP, TPSA and log P of a chemical may play an important role in oral and/or dermal uptake (Potts and Guy 1992, Pugh et al. 2000, Lipinski et al. 2001, Magnusson et al. 2004).

By linking the DA values from the publically available SCCS opinions to the physicochemical properties MW, MP, TPSA and log P, we have shown that rules can be extracted to identify compounds suspected to have a low DA and which may be associated with a low dermal bioavailability.

According to this study the rule set showed a sensitivity of 100% and a specificity of 20%. After setting new boundary criteria and applying more flexible rules the performance of the rule set was optimised. The sensitivity of the predictions remained 100%, implying that compounds with a high DA are always predicted as such and the specificity was increased to 58%, without compromising the sensitivity. It is indeed preferable not to identify compounds with a high DA as having a low DA. So in case a compound triggers none or only one of the following alerts: MW < 180 Da, log P $\geqslant$ 0.3, MP < 100°C or TPSA < 40 Å^2, it is likely to have a low DA and thus a low dermal bioavailability. The

267 presented rule set offered the best consensus between specificity and sensitivity. Adding more
268 criteria to classify a compound as high dermal absorption, did not lead to a better prediction.
269 It should be noted that this study comprises a limited set of compounds. Furthermore, several
270 difficulties were encountered in modelling these DA data. The data have been collated from the
271 results of different assays using inconsistent methodologies in terms of species, exposure times,
272 concentrations, matrices, detection methods etc. Also, it must be noted that the data set analysed
273 here is skewed very much towards low DA values and that the same rules may not apply when
274 investigating compounds from different chemical domains. Also, the possibility of bioaccumulation is
275 not taken into account in this study. Nonetheless, this pragmatic approach shows that when
276 physicochemical evidence suggests that a cosmetic ingredient has a low DA and thus low dermal
277 bioavailability, it might be worthwhile to further investigate this by performing more extensive *in*
278 *vitro* DA studies to get more reliable mean values and to confirm the very low DA (*i.e.* testing
279 different concentrations, using relevant excipients, increased sample size…). Although the data are
280 skewed towards low DA, in many cases the DA value used in the SCCS safety dossiers is still over-
281 estimated and more extensive *in vitro* DA studies might enforce the reliability of the obtained results.
282 Especially when taking into account that two standard deviations are added to the mean DA value
283 when the variability between the different measurements is high or when the DA studies have not
284 been carried out under ideal test conditions.
285
286 To add further to the weight of evidence, existing computational tools could be used to predict oral
287 bioavailability (Moda et al. 2007; Kumar et al. 2011). In case oral and dermal bioavailability are both
288 low, it would strengthen the safety assessor's reasoning to omit the need to calculate the MoS,
289 making at least for this type of ingredients *in vivo* repeated dose toxicity studies redundant and to
290 focus on local toxicity (skin sensitisation and irritation) and mutagenicity/genotoxicity test results.
291 Since most of the existing computational tools have been developed for pharmaceuticals, evidence
292 should be provided for their applicability in the cosmetic sector.

To notice for the future is the possibility that when substantial evidence of low bioavailability is provided, the internal threshold of toxicological concern (TTC) concept might be applied. This probabilistic approach is used to identify human exposure thresholds below which the risk of toxicological concern is low by taking into account oral/dermal absorption of the compound (internal exposure) rather than external exposure (Partosch et al. 2015). For completeness, decisions relating to internal exposure following oral/dermal administration should include considerations of metabolism when one wants to omit the determination of the NOA(E)L, since it will then be important to consider the possibility of metabolic activation. Several *in vitro* and *in silico* models are available for predicting metabolism following oral exposure and there is increasing interest in the area of skin metabolism for which models are currently being developed (as reviewed recently by Dumont et al 2015). Though in the cosmetic sector the TTC concept has been accepted for the safety assessment of impurities for which the identity is known but toxicity data are lacking (Kroes et al. 2007; SCCS/1564/15), more evidence is still needed to prove the applicability of the internal TTC for cosmetic ingredients.

Acknowledgements

This work was supported by Vrije Universiteit Brussel.

References

2009/1223/EC. Regulation (EC) No 1223/2009 of the European Parliament and of the Council of 30 November 2009 on cosmetic products. *Official Journal* 2009; *L 342:*59.

Chu KA, Yalkowsky SH. An interesting relationship between drug absorption and melting point. *International Journal of Pharmaceutics* 2009; 373(1-2):24-40. **DOI**: 10.1016/j.ijpharm.2009.01.026

Dumont C, Prieto P, Asturiol D, Worth A. Review of the availability of *in vitro* and *in silico* methods for assessing dermal bioavailability. *Applied in Vitro Toxicology* 2015; 1(2): 147-164. **DOI**: 10.1089/aivt.2015.0003

Kroes R, Renwick AG, Feron V, Galli CL, Gibney M, Greim H, Guy RH, Lhuguenot JC, van de Sandt JJM. Application of the threshold of toxicological concern (TTC) to the safety evaluation of cosmetic ingredients. *Food and Chemical Toxicology* 2007; 45: 2533–62. **DOI**: 10.1016/j.fct.2007.06.021

Kumar R, Sharma A, Kumar Varadwaj P. A prediction model for oral bioavailability of drugs using physicochemical properties by support vector machine. *Journal of Natural Science, Biology and Medicine* 2011; 2(2): 168-73. **DOI**: 10.4103/0976-9668.92325

Lipinski CA, Lombardo F, Dominy BW, Feeney PJ. Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. *Advanced Drug Delivery Reviews* 2001; 46(1-3):3-26. **DOI**: 10.1016/S0169-409X(00)00129-0

Magnusson BM, Anissimov YG, Cross SE, Roberts MS. Molecular size as the main determinant of solute maximum flux across the skin. *Journal of Investigative Dermatology* 2004; 122(4):993-9. **DOI**: 10.1111/j.0022-202X.2004.22413.x

Moda TL, Montanari CA, Andricopulo AD. Hologram QSAR model for the prediction of human oral bioavailability. *Bioorganic and Medicinal Chemistry* 2007; 15(24): 7738-45. **DOI**: 10.1016/j.bmc.2007.08.060

Newby D, Freitas AA, Ghafourian T. Decision trees to characterise the roles of permeability and solubility on the prediction of oral absorption. *European Journal of Medicinal Chemistry* 2015; 90:751-65. **DOI**: 10.1016/j.ejmech.2014.12.006

Partosch F, Mielke H, Stahlmann R, Kleuser B, Barlow S, Gundert-Remy U. Internal threshold of toxicological concern values: enabling route-to-route extrapolation. *Archives of Toxicology* 2015; 89:941-48. **DOI**: 10.1007/s00204-014-1287-6

Potts RO, Guy RH. Predicting skin permeability. *Pharmaceutical Research* 1992; 9:663-9.

Pugh WJ, Degim IT, Hadgraft J. Epidermal permeability-penetrant structure relationships: 4, QSAR of permeant diffusion across human stratum corneum in terms of molecular weight, H-bonding and electronic charge. *International Journal of Pharmaceutics* 2000; 197(1-2):203-11.

SCCS/1501/12. The SCCS's notes of guidance for the testing of cosmetic substances and their safety evaluation – 9th revision. 2015 Available through: http://ec.europa.eu/health/scientific_committees/consumer_safety/docs/sccs_o_190.pdf, consulted November 2015.

Steinmetz FP, Madden, JC, Cronin, MTD. Data Quality in the Human and Environmental Health Sciences: Using Statistical Confidence Scoring to Improve QSAR/QSPR Modeling. *Journal of Chemical Information and Modeling* 2015; **DOI:** 10.1021/acs.jcim.5b00294

Veber DF, Johnson SR, Cheng H-Y, Smith BR, Ward KW, Kopple, KD. Molecular properties that influence the oral bioavailability of drug candidates. *Journal of Medicinal Chemistry* 2002; 45(12): 2615-2623. **DOI**: 10.1021/jm020017n