DENTAL MATERIALS XXX (XXXX) XXX–XXX

Available online at www.sciencedirect.com

ScienceDirect

journal homepage: www.elsevier.com/locate/dental

# First multi-target QSAR model for predicting the cytotoxicity of acrylic acid-based dental monomers

*Amit Kumar Halder*[a,b], *António H.S. Delgado*[c,d,*], *M. Natália D.S. Cordeiro*[a,**]

[a] LAQV@REQUIMTE/Department of Chemistry and Biochemistry, Faculty of Sciences, University of Porto, Porto, Portugal
[b] Dr. B. C. Roy College of Pharmacy and Allied Health Sciences, Dr. Meghnad Saha Sarani, Bidhannagar, Durgapur 713212, West Bengal, India
[c] Division of Biomaterials and Tissue Engineering, UCL Eastman Dental Institute, London, UK
[d] Centro de Investigação Interdisciplinar Egas Moniz (CiiEM), Monte de Caparica, Portugal

ARTICLE INFO

*Article history:*
Received 30 April 2021
Received in revised form
17 November 2021
Accepted 8 December 2021
Available online xxxx

*Keywords:*
Dental monomers
Cytotoxicity
Multi-target QSAR models
Machine learning
Similarity searching

ABSTRACT

*Objective:* Acrylic acid derivatives are frequently used as dental monomers and their cytotoxicity towards various cell lines is well documented. This study aims to probe the structural and physicochemical attributes responsible for higher toxicity of dental monomers, using quantitative structure-activity relationships (QSAR) modeling approaches.

*Methods:* A regression-based linear single-target QSAR (st-QSAR) model was developed with a comparatively small dataset containing 39 compounds, the cytotoxicity of which has been assessed over the Hela S3 cell line. By contrast, a classification-based multi-target QSAR model was developed with 138 compounds, the cytotoxicity of which has been reported against 18 different cell lines. Both models were set up following rigorous validation protocols confirming their statistical significance and robustness.

*Results:* The performance of the linear mt-QSAR model, developed with various feature selection and post-selection similarity searching-based schemes, superseded that of all non-linear models produced with six machine learning methods by hyperparameter optimization. The final derived st-QSAR and mt-QSAR linear models are shown to be highly predictive, as well as revealing the crucial structural and physicochemical factors responsible for higher cytotoxicity of the dental monomers.

*Significance:* This study is the first attempt on unveiling the cytotoxicity of dental monomers over several cell lines by means of a single multi-target QSAR model. Further, such a model is ready to get widespread applicability in the screening of new monomers, judging from its almost accurate predictions over diverse experimental assay conditions.



* Correspondence to: Biomaterials and Tissue Engineering Department - Royal Free Hospital, UCL Medical School, Rowland Hill Street, Hampstead, NW3 2PF, London, UK.
** Correspondence to: LAQV@REQUIMTE/Department of Chemistry and Biochemistry, Faculty of Sciences, University of Porto, Rua do Campo Alegre, s/n, 4169–007 Porto, Portugal.
E-mail addresses: antonio.delgado.17@ucl.ac.uk (A.H.S. Delgado), ncordeir@fc.up.pt (M.N.D.S. Cordeiro).

https://doi.org/10.1016/j.dental.2021.12.014
0109-5641/

## 1. Introduction

Polymers are a family of materials widely used in all dental specialties in a variety of clinical scenarios. Dental polymers are made of methacrylate-based monomers which are generally monomethacrylates or dimethacrylates [1,2]. These can be bisphenol-A glycidil dimethacrylate (Bis-GMA), urethane dimethacrylate (UDMA), triethylene glycol dimethacrylate (TEGDMA), 2-hydroxyethyl methacrylate (HEMA) or other more recent functionalized monomers [1,3]. In most cases, such materials are in direct contact with oral tissues, and biocompatibility is imperative for clinical use [1,4]. Resin composites and dental adhesives used for direct restorations that feature acrylic-acid based monomers are placed in contact with the sensitive dentin-pulp complex [5]. Monomers of these restorative materials may diffuse through the dentinal tubules and elicit toxic effects to pulp cells, compromising the vitality of the tooth [1,6]. Biological effects such as estrogenicity and genotoxicity of Bis-GMA, or mutagenicity of other monomers have also been linked to dental materials and are well-documented in past studies [1, 7–10].

Dental materials that include methacrylate-based monomers undergo a free-radical addition polymerization reaction which results in incomplete monomer conversion, originating residual unreacted monomers [1,11]. The latter pose toxicity risks as they can leach into surrounding aqueous phases and are then able to enter the organism [10,12]. Monomers are able to elute from the set polymer either due to incomplete polymerization or natural degradation of the matrix [13]. The resin phase in composites suffers degradation in the oral environment with aging, and phenomenon such as hydrolysis leads to breakdown of the organic matrix and release of monomers [1, 14, 15]. Such biodegradation of resin composites may have the potential to cause oral mucosa irritation, allergic reactions or even loss of bone [16]. Degradation and subsequent leaching both depend upon physicochemical properties such as different functional groups, molecular size of monomers and chemical structure [1,17].

Quality control that includes biocompatibility studies is therefore extremely important in dental materials research, and cytotoxicity testing is an important assay which measures cell response and death to novel material formulations [18]. Cytotoxic effects should be preventable and minimized, and that is the goal when materials are being developed, especially concerning fields such as operative dentistry and endodontics [18,19]. Cytotoxicity related to dentistry has been tested with different assays such as 3-(4,5-dimethylthiazol-2-yl)-2,5-diphenyl tetrazolium bromide (MTT), live-dead assay, modified sulforhodamine B assay (SRB), lactate dehydrogenase (LDH) and in several cell lines which include mouse odontoblasts (MDPC-23), mouse fibroblasts (3T3 or L929), human fibroblasts (FP5), rat pulp cells (RPC-C2A), Hela S3 or peripheral blood mononuclear cells [10, 15, 20–22]. However, systematic reviews turn out complaining about the lack of methodological standardization in the primary studies that evaluated toxicity, aside from the fact that cell studies are laborious and time consuming [22].

Nowadays, alternative *in silico*-based modeling approaches, like the application of Quantitative Structure-Activity Relationships (QSAR), offer a viable option in order to minimize bench work and for more effectively screening novel molecules during materials' development [23,24]. Indeed, QSAR models had previously been developed with success to predict material characteristics such as the lipophilicity of dental monomers and even their mutagenicity [25,26]. Yet an extension of those models to include the cytotoxicity of dental monomers using different cell lines and endpoints has still not been attempted. Thus, the aim of this study is to develop validated QSAR models able to predict the cytotoxicity of acrylic acid-based monomers frequently used in the dental field and, additionally, as a screening-based scheme for guiding novel formulations.

In 1997, Yoshii reported the cytotoxicity data of 39 acrylic dental monomers against Hela S3 cell lines [20]. However, the cytotoxicity of dental monomers has been reported also against various cell lines, using different experimental assay conditions. Interestingly, it has been found that the latter varied considerably depending on the type of cell line and measure of effects. Such fact justified the application of *in silico*-based approaches with a wide range of applicability, such as multi-target QSAR (mt-QSAR) modeling techniques. QSAR modeling is today an established data-driven method with proven efficacy in screening and evaluating compounds for medical use [27]. Compared to large screening methods for novel biological and chemical materials, such as high-throughput screening (HTS), QSAR modeling allows larger amounts of compounds to be included in the screening process while maintaining a lower cost [27]. On the other hand, large datasets obtained from HTS also fostered an urgent need for big data analysis through advanced QSAR modeling [24]. QSAR models have been around for more than 60 years but have witnessed significant improvements ever since. QSAR modeling is encouraged by renowned regulatory agencies, such as the Food and Drug Administration (FDA) or the Registration, Evaluation, Authorization and Restriction of Chemicals (REACH) [28], and it provides essential means for filling the knowledge gaps. Furthermore, to guide and regularize the best practices across different countries, while scientifically validating these models, regulatory principles for QSAR models have been set out by the Organization for Economic Co-Operation and Development (OECD). The OECD, as thoroughly reviewed by Cumming et al. in 2013 [29], requires the model to have a defined end point, an unambiguous algorithm with a defined domain of applicability, goodness of fit, robustness and predictive ability, as well as a mechanistic interpretation if possible. Compared to conventional single target QSAR modeling, multi-target QSAR modeling based on the Box-Jenkins moving average approach is a comparatively novel technique that is able to simultaneously predict the cytotoxicity response pertaining to multiple experimental assay conditions [30–33], and at the same time, this technique follows the OECD guidelines to ensure the reliability and transparency of the developed mt-QSAR models. The scope of the current work is not only to limited to the development of QSAR models to predict cytotoxicity of acrylic acid based dental monomers through single and multi-target QSAR modeling techniques, but it also proposes a new technique named 'Post selection similarity search-based modification' (PS3M) that was found to be highly effective in finding highly predictive linear regression as well as classification models,

and therefore these kind of techniques may be utilized in the future for developing linear QSAR models. Similarly, this work reports for the first time an *in house* tool named PS3M (to be accessed from https://github.com/ncordeirfcup/PS3M), which is now available in the public domain for its application.

## 2. Materials and methods

### 2.1. Datasets

The data set compiled for this study includes 138 cytotoxicity data-points (MD1-MD138) of 58 acrylic acid based dental monomers that were retrieved from a comprehensive literature search (please refer to the Supplementary Material - Table S1). In this dataset, the cytotoxicity of these monomers was assayed against 18 different biological targets/cell lines and their cytotoxicity probed by five types of measure of effects, namely: $ED_{50}$, $TC_{50}$, $IC_{50}$, $LC_{50}$ and $EC_{50}$. Only data from studies which carried out a MTT assay were included in this study. This dataset was used for the development of a multi-target classification QSAR (mt-QSAR) model and is henceforth referred as multi-target dataset. Now in the same dataset, $IC_{50}$ values of 39 acrylic acid based dental monomers were reported only against Hela S3 cell lines [20]. Therefore, this 39 compound-dataset (MD1-MD39; Table S1), which hereafter will be referred as single-target dataset, was used for building the single target QSAR (st-QSAR) model.

Naturally, single- and multi-target datasets have distinct characteristics and therefore st- and mt-QSAR models derived with them also serve different purposes. The scope of conventional st-QSAR modeling is limited to only one experimental assay condition (*i.e.*, one biological target and type of end-point), whereas mt-QSAR modeling affords incorporation of as many as possible experimental assay conditions. As such, mt-QSAR not only increases the overall chemobiological space of the resultant *in silico* models but also improves their overall applicability. The current work is of no exception. The application and reliability of the st-QSAR model will be limited to one biological target (*i.e.*, Hela S3) and measure of toxic effect (*i.e.*, $IC_{50}$). On the other hand, as the derived mt-QSAR model relates to a wide range of experimental conditions [30], its overall applicability should be higher than that of the st-QSAR model. The set-up of the mt-QSAR model is based on a dataset, in which the same compound was found to have different responses against different cell lines. As an example, 2-hydroxymethacrylate showed relatively high cytotoxicity against cell lines such as JTC-12 ($IC_{50}$ = 1.692 mM), 3T3 ($ED_{50}$ = 1.77 mM), HPLF ($ED_{50}$ = 1.87 mM), *etc.*, whereas the same compound depicted relatively low cytotoxicity against other cell lines such as THP-1 ($TC_{50}$ = 11 mM), 3T3 ($TC_{50}$ = 13.46 mM) and PBMC ($TC_{50} > 300$ mM). Highly predictive mt-QSAR modeling thus attempts to disclose these complicated relations between the chemicals and biological targets. One should also note here that the number of datapoints against each experimental condition vary to a considerable extent in the multi-target dataset. It is thus not possible to build a statistically reliable st-QSAR model for each of the latter conditions. Owing to this, the major purpose of the present work is to set up linear and non-linear mt-QSAR models based on such multi-target dataset. Nevertheless, the linear st-QSAR model is being reported also here and compared with the respective mt-QSAR model to check if any common structural attribute might be detected for the targeted monomers.

### 2.2. Molecular structures and descriptors

Molecular structures of the dataset compounds were firstly inputted from the SMILES (simplified molecular input line entry specification) codes retrieved from Pubchem (https://pubchem.ncbi.nlm.nih.gov/) and subsequently converted to sdf format (without adding explicit hydrogen atoms), using the MarvinView tool (MarvinView. Version 18.18.0; ChemAxon: Budapest, Hungary, 2010, https://chemaxon.com/products/marvin). The latter were then standardized with the help of the Standardizer tool of Chemaxon using the following options: (a) add explicit hydrogen atoms, (b) aromatize, (c) clean 2D, (d) clean 3D, (e) neutralize and (f) strip salts (Standardizer. Version 15.9.14.0 Software; ChemAxon: Budapest, Hungary, 2010). Different sets of 2D- and 3D-molecular descriptors were calculated for these dataset compound structures using the newly launched alvaDesc software (https://www.alvascience.com/alvadesc/) with the help of OCHEM platform [34]. In OCHEM, structure optimization of the compounds was performed with Corina [35] for the calculation of 3D molecular descriptors.

### 2.3. Single target QSAR modeling

To begin with, the $IC_{50}$ (mM) values of the single-target dataset were log-transformed ($pIC_{50}$) for being of practical use in the following st-QSAR modeling. The dataset was then divided into three training set and test set combinations using an activity-based sorting scheme, which is especially useful for small datasets. In this data-division scheme, the dependent variable (*i.e.*, $pIC_{50}$) is first sorted out in descending manner and test data are subsequently collected with a starting point (the data-point from which data collection initiates) and an interval obtained from the percentage of data to be included in the test set. In the present work, we resorted to three starting points, *i.e.*,: 1, 2 and 3, whereas 20% of data was collected for each starting point to build three separate data-distributions. Subsequently, multiple linear regression (MLR) models were set up by applying the sequential forward selection (SFS) technique to select the independent $X$ – variables (molecular descriptors) with the highest impact on the cytotoxicity. The MLR models were developed separately for each of these data-distributions, using our *in-house* tool SFS-QSAR (available at https://github.com/ncordeirfcup/SFS-QSAR-tool) that resorts to the 'Feature Selector' module of the library *Mlxtend* (http://rasbt.github.io/mlxtend/) [36]. To remove highly intercorrelated and less-variant descriptors, a correlation cut-off of 0.95 and variance cut-off of 0.001 were set in SFS-QSAR. During model development, four scoring functions of the 'Sequential Feature Selector' module were employed for feature selection, namely: the determination coefficient ($R^2$), the negative mean absolute error (*NMAE*), the negative mean Poisson deviance (*NMPD*), and the negative mean gamma deviance (*NMGD*). No cross-validation was performed during feature selection.

Comparison of the quality of the different st-QSAR regression models was carried out by means of various internal and external diagnostic statistical tools. Measures of their internal statistical significance were estimated by standard validation parameters such as $R^2$, the leave one out cross-validation $R^2$ ($Q^2_{LOO}$), the mean absolute error (*MAE*), along with two scaled $r_m^2$ metrics, *i.e.*: $r^2_{mLOO}$ and $\Delta r_m^2$ (The latter metrics are calculated based on the correlation between the observed and (LOO)-predicted values with and without intercept for the least squares regression lines [37].). Similarly, the predictive ability of the models was assessed by the following external validation parameters: $R^2_{Pred}$, $r^2_{mtest}$ (scaled) and $\Delta r^2_{mtest}$ (scaled) [38–40]. To enhance the statistical performance of the most predictive MLR model found, the latter was further subjected to the technique that will be called hereafter 'Post selection similarity search-based modification' technique (see description below).

### 2.4. *Multi-target QSAR modeling*

Amongst others, the mt-QSAR model development technique based on the Box-Jenkins moving average approach has proven to be especially robust for coping with heterogeneous data, mainly due to its versatility as well as simplicity [41–45]. Therefore, in the present work, the Box-Jenkins approach was applied to develop mt-QSAR classification-based models for predicting the cytotoxicity of acrylic acid monomers against the several experimental assay conditions. Details of the latter approach have been largely described in the past [33,46], so the focus here will only be towards the most important aspects related to this study.

The experimental conditions ($c_j$) under which the chemicals understudy have been tested are better expressed as an ontology of the form $c_j \rightarrow$ (*bt*, *me*). In such an ontology, the first element, *bt*, refers to the cell lines against which the cytotoxicity of the monomers has been assayed, whereas the second one, *me*, simply stands for the different type of measures of such endpoint response. Furthermore, since our aim is to set up a classification-based mt-QSAR model, a cut-off value was chosen for discriminating toxic monomers from non-toxic ones. In this study, monomers with a cytotoxicity less than 2 mM (irrespectively of *me*) were taken as toxic whereas the remaining ones as non-toxic. In so doing, 86 data-points were categorized as toxic ($IA_i(c_j) = +1$) and 52 data-points as non-toxic ($IA_i(c_j) = -1$).

To further discriminate the monomers' cytotoxicity when the different elements of $c_j$ are varied, the initial molecular descriptors ($D_i$ calculated with alvaDesc) should be converted to new modified descriptors following the Box-Jenkins moving average approach, *i.e.*, as follows:

$$\Delta(D_i)c_j = D_i - avg(D_i)c_j \quad (1)$$

where $avg(D_i)c_j$ is the arithmetic mean of descriptors $D_i$ for the number of toxic monomers in the dataset assayed under the same element of the experimental condition (ontology) $c_j$ [33].

However, Eq. (1) stands for the simplest form of defining the modified descriptors and recently, other schemes have been proposed. Alternatively, for example, normalized modified descriptors can be calculated by considering the difference between the maximum ($D_{i\ max}$) and minimum ($D_{i\ min}$) values of input descriptors [43]:

$$\Delta(D_i)c_j = \frac{D_i - avg(D_i)c_j}{D_{i\,max} - D_{i\,min}} \quad (2)$$

Likewise, normalization can be performed also regarding the experimental elements of $c_j$ [41], *i.e.*, as follows:

$$\Delta(D_i)c_j = \frac{D_i - avg(D_i)c_j}{(D_{i\,max} - D_{i\,min})\,p(c_j)_c} \quad (3)$$

where $p(c_j)_c$ stands for the *a priori* probability of finding datapoints tested under a specific assay condition $c_j$.

In this study, normalized modified descriptors computed according to Eqs. (2) and (3) were used to set up both the linear and non-linear mt-QSAR classification models. In fact, normalization of those descriptors was required due to the large structural variations in the dataset, as can be observed from the distribution of molecular weights (MW) and lipophilicity (MLOGP2) of the present monomers (see Fig. 1). One should also notice that the number of occurrences of experimental elements in the dataset vary to a considerable extent, making $p(c_j)_c$ important and justifying the selection of the other type of normalized modified descriptors (Eq. (3)). All the calculations were carried out using our recent developed open source python-based toolkit *QSAR-Co-X* [46] (available to download at https://github.com/ncordeirfcup/QSAR-Co-X).

Similar to the single-target dataset, the multi-target dataset including the modified descriptors were divided into different data-distributions for the sake of external validation of the models. Two methods were used to divide the multi-target dataset into a training and a validation set, namely: (i) random division and (ii) *k*-means cluster analysis (*k*MCA) [42,47]. In both cases, the test set size was chosen to be 20% of the whole data set. In the *k*MCA-based division approach, five clusters were initially formed using the response variable and dependent parameters. From each of these clusters, 20% of the data was randomly selected as the test set. The training set was used for calculation of the $avg(D_i)c_j$, $D_{i\ max}$, $D_{i\ min}$ and $p(c_j)_c$ values needed to obtain the modified descriptors of such set – *i.e.*,: the values of $\Delta(D_i)c_j$ using Eqs. (2) and (3) separately, and these were then used for calculating the modified descriptors of the validation set. The training set was further divided into a sub-training set (80%) and a test set (20%). The models were first developed solely using the sub-training set and the predictive accuracy of the models were first tested with the test set. The external predictivity of the most predictive model was then tested with the validation set, which may be considered as a 'true validation set' because it neither participated in the descriptor calculation method nor in the model development.

To set up the linear mt-QSAR classification models, we opted for the linear discriminant analysis (LDA) technique [48], along with two different feature selection algorithms, namely: (i) fast stepwise (FS) and (ii) sequential forward selection (SFS). In FS-LDA, the descriptors were selected on the basis of the *p*-value, whereas in SFS-LDA those were selected using the 'Accuracy' as scoring function. For both cases, only a maximum number of five descriptors ($X$ – variables) was allowed to be included, and all descriptors with inter-correlation higher than 0.95 and variance less than 0.001 were removed before model development.

Goodness-of-fit of derived the linear models were initially checked by standard statistics such as the Wilk's λ parameter

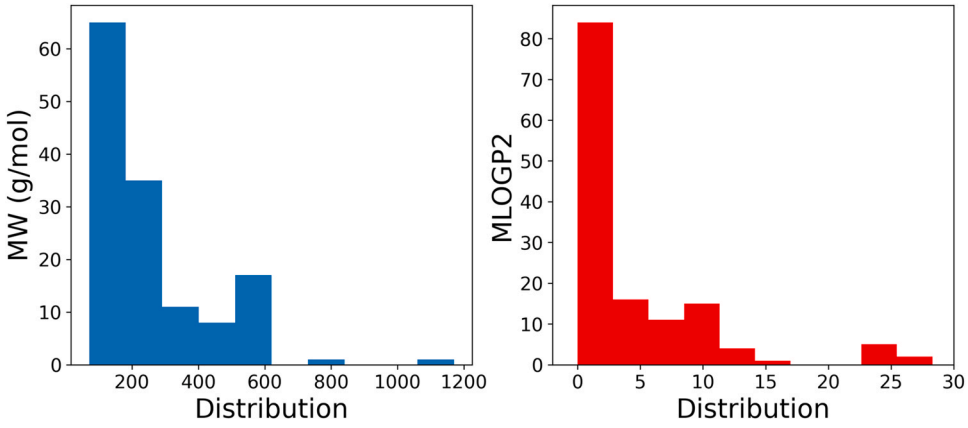


**Fig. 1 – Histogram plots displaying the distribution of molecular weights (MW) and lipophilicity (MLOGP2) for the monomers of the multi-target dataset.**

[49], the Fisher ratio (*F*), and the corresponding *p* and Chi square ($\chi^2$) values. Internal and external predictivity of the models was assessed by evaluating metrics such as sensitivity (*Sn*), specificity (*Sp*), accuracy (*Acc*), and the F1-score (see Eqs. (4)–(7) below), as well as by both the Matthews correlation coefficient (MCC) and the area under receiver operating characteristic (AUROC) [50–52]. Alike the st-QSAR linear modeling, the PS3M technique was also implemented for attempting to improve the statistical quality of the most predictive linear mt-QSAR model.

$$Sn = \frac{TP}{(TP + FN)} \tag{4}$$

$$Sp = \frac{TN}{(TN + FP)} \tag{5}$$

$$Acc = \frac{(TP + TN)}{(TP + TN + FP + FN)} \tag{6}$$

$$F1 - score = \frac{2TP}{(2TP + FP + FN)} \tag{7}$$

where TP, TN, FP and FN stand for the number of true positive, true negative, false positive and false negative samples, respectively.

Regarding the non-linear mt-QSAR classification models, these were set up by resorting to six machine learning (ML) classifier techniques [53] using the QSAR-Co-X toolkit [43], *i.e.*,: (i) *k*-Nearest Neighborhood (*k*NN) [54], (ii) Bernoulli Naïve Bayes (NB) [55], (iii) Support Vector Classifier (SVC) [56], (iv) Random Forests (RF) [57], (v) Gradient Boosting (GB) [58] and (vi) Multi-Layer Perception (MLP) [59]. For each technique, a hyperparameter tuning was applied to select the most suitable model development parameters, based on a 5-fold cross-validation (CV) scheme. Details about the tuned parameters are given in Table S2 of the Supplementary Material.

Internal predictivity of the derived non-linear models was checked with the help of 5-fold CV performed with the sub-training set using the optimized ML parameters. In what regards their external predictivity, that was first assessed with the test set and finally with the validation set. Identical metrics as before were calculated for both the sub-training, test and validation sets to assess the internal and external predictivity of these models – *i.e.*,: *Sn*, *Sp*, *Acc*, F1-score, MCC and AUROC values [50–52].

### 2.5. *Y-based randomization and applicability domain*

Model validation proceeds usually *via* the Y-based randomization technique – *i.e.*, by response randomization testing. In the current work, the Y-randomization technique was applied to the regression-based st-QSAR linear model by randomizing 1000 times its responses and then, the metric $cR_P^2$ was calculated to check the uniqueness of the model [60]. In a similar way, the *Yc*-based randomization technique (*i.e.*,: by response along with ontology randomization testing) was applied to the linear mt-QSAR model as described earlier [43]. In this case, after *Yc*-based randomization, both the average accuracy and the average Wilk's λ parameter values were then compared to those obtained with the original mt-QSAR model to establish the uniqueness of the latter.

Another goal in any QSAR modeling is to establish the applicability domain (AD) of the final derived model, that is, the range within which it makes predictions with a given reliability. For such purpose, the AD of the st-QSAR linear model was checked by plotting the standardized residuals *vs.* the leverage values for each monomer understudy. From this plot, the so-called William's plot [61,62], outliers are identified by checking whether these are outside a squared area (AD) within ± 3 times the standardized residuals and the leverage threshold $h^*$ ($h^* = 3*(p+1)/n$, with *p* the number of descriptors in the model and *n* the total number of training set samples). In the case of the mt-QSAR models, the simpler standardization approach suggested by Roy et al. [63] was adopted to establish their applicability domain.

### 2.6. *Post-selection similarity search-based modification scheme*

Linear interpretable models may be developed using various feature selection algorithms. Yet none of the latter can be considered the best one and in reality, numerous possible linear models may coexist with almost equal predictivity [64]. As referred to above, both st-QSAR and mt-QSAR linear

models were developed using forward feature selection algorithms based on different scoring functions and statistics. Nonetheless, the resulting models depend on multiple factors such as data-distribution, descriptor pre-treatment as well as scoring functions. That is, despite having high statistical quality, the final linear model selected may not be the most predictive and actually, it may just be considered as a reference model. Models with a similar statistical quality can thus be compared to such reference model to check whether they provide better predictive power or not. These models in turn can easily be obtained from the reference model by replacing its descriptors with others closer in terms of a specific statistical parameter, such as the Euclidean distance.

In the 'Post-selection similarity search-based modification' (short form PS3M), the reference linear model is trained with all descriptors of the modeling data-matrix and each of its descriptors is then used to find the $m$ number (user specific) of descriptors that have minimum Euclidean distance (ED) from it. ED, the distance between two descriptors (D1 and D2) is simply calculated as follows:

$$d(D_1, D_2) = \sum_{i=1}^{n} (D1_i - D2_i)^2 \quad (8)$$

where $n$ is the number of data-points.

As such, if the reference model contains $p$ number of descriptors, after removing from it descriptors with ED ~ 0, we may expect to get ($m^*p$) number of alternative models that could be referred as '*similar*' linear models. The 'Post-selection similarity search-based modification' assumes that a more predictive linear model may be found from these similar alternative models. Evidently, it should not be wise to select a large value for $m$ because it would reduce the similarity of the alternative models with the reference model and may give rise to underfitted/overfitted models. In this work, we therefore fixed $m$ as 10 for both regression and classification models. If any better model was found, it was then used as a next reference model for a second run and these runs were continued until no better model was found. For selection of the better regression model, we used the average $r_m^2$ statistics (*i.e.*,: average values of $r_{mLOO}^2$ and $r_{mtest}^2$) whereas for the classification model, we used the average MCC value obtained from different data-distributions. All PS3M refinements were carried out with our *in-house* developed tool, which is available at https://github.com/ncordeirfcup/PS3M.

## 3. Results and discussion

### 3.1. Linear st-QSAR model

Three different training set (80%)-test set (20%) data-distributions, namely: AS1 (start point: 1), AS2 (start point: 2) and AS3 (start point: 3), were obtained by the activity sorting method [65]. Subsequently four scoring functions were used for feature selection to obtain twelve st-QSAR linear models. The statistical results of these models (SL1-SL12), each generated with five descriptors, are summarized in Table S3 of the Supplementary Material. As seen, the most predictive model SL7 was obtained with the data-distribution AS2, for which the negative mean Poisson deviance (NMPD) was used as scoring function. Judging from $Q_{LOO}^2$ of 0.826 and $R_{Pred}^2$ of 0.759, SL7 may be regarded as a statistically robust linear QSAR model. The model was then subjected to the PS3M technique and after a single run, one similar model (SL7a) was found with improved internal and external predictivity results. No better model could be found by further PS3M and therefore this model was accepted as the final st-QSAR linear model. Table 1 shows in detail the statistical results obtained for both these linear models.

As seen in Table 1, the PS3M technique returned the final model SL7a by simply replacing Mor04v descriptor with Mor31u (ED = 3.143). As anticipated, the latter two descriptors are highly correlated (Pearson $R$ = 0.86). Even so, the replacement give rise to a model with improved statistical parameters since model SL7a is able to account for 86.3% and 84.5% of the explained and predicted variance of the data. The values obtained for the $r_m^2$ metrics (*i.e.*,: $r_{mLOO}^2 = 0.780$ and $\Delta r_{mLOO}^2 = 0.115$) are also indicative of its good internal predictivity. Further, the SL7.1 model depicts satisfactory external predictivity, judging from the values attained for $R_{pred}^2$, $r_{mtest}^2$, and $\Delta r_{mtest}^2$ (see Table 1). Additionally, the model lacks high inter-collinearity among its descriptors as the maximum $R^2$ between any two descriptors was found to be particularly small (= 0.075). Moreover, the application of the Y-based randomization test led to a value of 0.779 for $cR_P^2$, confirming that the model was not generated by chance.

Fig. 2 shows the observed *vs*. predicted cytotoxicity and the Williams plot for the final SL7a model. As can be seen in the Williams plot, all the monomers are inside of the AD squared area, save for monomer MD38 (see Table 1 of the Supplementary Material). Even so, since the latter has a leverage greater than $h^*$ but shows a standard residual value within the limits, it should only be considered as an influential monomer rather than an outlier. The meaning of the descriptors appearing in this final model is given in Table 2 along with their relative importance according to the corresponding regression coefficients (see Table 1). Among its five descriptors, VE1sign_B(v) was found to be the most influential one followed by GATS5m, C-005, Mor31u and H-052. As such, one can assume that higher atomic van der Waals volumes/atomic masses/presence of certain chemical fragments could induce an increase/decrease/decrease, respectively, of the cytotoxicity for the present dental monomers against the Hela S3 cell line.

### 3.2. Linear mt-QSAR models

Moving on now to the mt-QSAR modeling of the dental monomers' cytotoxicity under different experimental conditions, the major goal of the present work. Following the strategy outlined before, we begun by developing twelve linear mt-QSAR models using multiple data-distributions. Subsequently, the most predictive model was selected based on both the goodness of fit and predictivity. A summary of the statistical results for these twelve models (models ML1-ML12) is given in Table S4 of the Supplementary Material. As can be seen, the predictive accuracy of ML1 superseded that of the other models and at the same time, the attained Wilk's $\lambda$ value (= 0.352) reinforces its statistical significance. The latter is also confirmed by the classification results; the

**Table 1 – Best derived st-QSAR linear models along with the MLR statistical results.**

| Model | St-QSAR model | Statistical results |
|---|---|---|
| **SL7** | $pIC_{50}$ = +5.124 (±0.385) + 2.278 (±0.651) VE1sign_B( v) − 2.000 (±0.338) GATS5m +0.383 (±0.072) Mor04v − 0.952 (±0.119) C − 005 − 0.105 (±0.040) H − 052 | $N_{training}$ = 31, $R^2$ = 0.877, $R^2_{Adj}$ = 0.853, F(25,5) = 35.71, $Q^2_{LOO}$ = 0.826, *MAE* = 0.198, $r^2_{mLOO}$ = 0.758, $\Delta r^2_{mLOO}$ = 0.092, $N_{test}$ = 8, $R^2_{Pred}$ = 0.759, $r^2_{mtest}$ = 0.574, $\Delta r^2_{mtest}$ = 0.210 |
| **SL7a**[a] | $pIC_{50}$ = +4.895 (±0.372) + 3.124 (±0.602) VE1sign_B(v) − 2.010 (±0.326) GATS5m +0.758 (±0.113) Mor31u − 1.007 (±0.115) C − 005 − 0.142 (±0.040) H − 052 | $N_{training}$ = 31, $R^2$ = 0.886, $R^2_{Adj}$ = 0.863, F(25,5) = 38.87, $Q^2_{LOO}$ = 0.845, *MAE* = 0.196, $r^2_{mLOO}$ = 0.780, $\Delta r^2_{mLOO}$ = 0.115, $N_{test}$ = 8, $R^2_{Pred}$ = 0.813, $r^2_{mtest}$ = 0.674, $\Delta r^2_{mtest}$ = 0.157 |

[a] Model refined from SL7 by applying the PS3M technique.

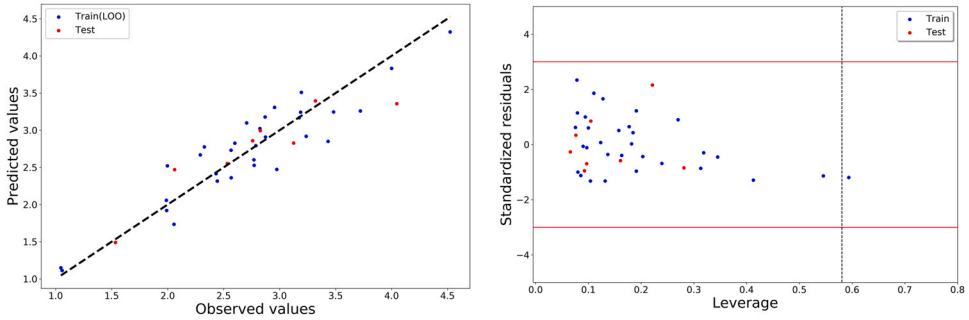


**Fig. 2 – Observed *vs.* predicted activity (left) and the Williams plot (right) obtained for model SL7a.**

**Table 2 – Meaning of the descriptors appearing in the final st-QSAR linear model SL7a.**

| Name | Type | Description[a] | Correlation with $pIC_{50}$ |
|---|---|---|---|
| **VE1sign_B(v)** | 2D matrix-based | Coefficient sum of the last eigenvector from Burden matrix weighted by van der Waals volume | Positive |
| **GATS5m** | 2D-Geary autocorrelation | Geary autocorrelation of lag 5 weighted by atomic mass | Negative |
| **C-005** | Atom-centered fragments | $CH_3X$ | Negative |
| **Mor31u** | 3D-MoRSE (Molecular Representation of Structures based on Electronic diffraction) | Unweighted 3D-Morse with scattering parameter s = 30 $Å^{-1}$ | Positive |
| **H-052** | Atom-centered fragments | Hydrogen atom attached to a $sp^3$ hybridized carbon atom with one heteroatom attached to next carbon | Negative |

[a] Refs [63,64].

model gives rise to an overall *ca.* 91% effective discrimination of both monomers from the sub-training and test sets, though slightly worse (~ 85%) for those from the external validation set. One should notice here that, removal of set monomer MD135 (see Table S1 of the Supplementary Material) from the validation set was carried out, since one of its experimental conditions (*i.e.*, bt: Balb/3T3 clone A31) was absent in the modeling dataset. Model ML1 is shown in Table S5 of the Supplementary Material, together with the corresponding LDA statistical parameters. Once more, the PS3M technique was applied to this model in an attempt to obtain a refined model with improved LDA statistics. Interestingly, the most predictive refined model ML1a was obtained after four steps as illustrated in Fig. 3.

The final refined mt-QSAR linear model ML1a is presented in Table 3, along with the statistical parameters of the LDA. Undoubtedly, the overall predictivity of this refined linear model is better than that of the original ML1 model, since it leads to better classification accuracy values for the sub-training, test and validation sets (of 94.3%, 90.9% and 88.9%, respectively). Similarly, the model gives rise to higher MCC scores of 0.882, 0.770 and 0.799 for the sub-training, test and validation sets, respectively. Even so, the goodness-of-fit of this model is slightly worse, judging from the higher attained

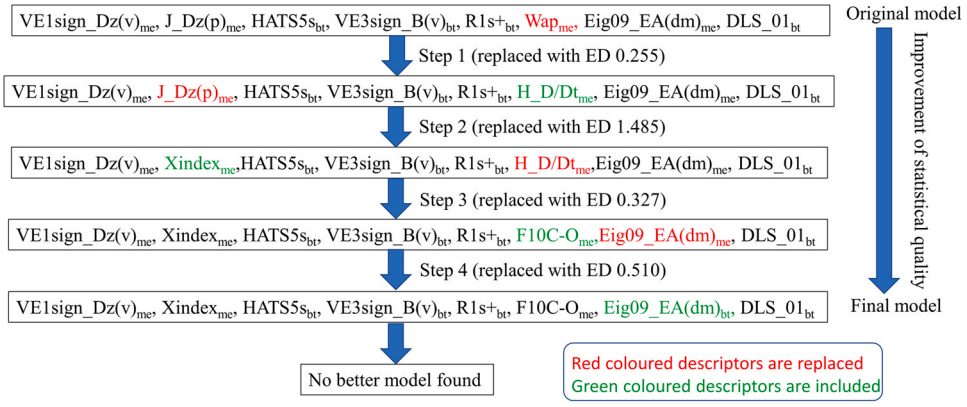


**Fig. 3 – Flowchart showing the stepwise modification of the original model ML1 (obtained from FS-LDA) to the final model ML1a (ED: Euclidean distance).**

Wilk's $\lambda$ value (= 0.356) compared to the one obtained for the ML1 model (= 0.352). Further, a 5-fold cross-validation performed with the sub-training set yielded an overall accuracy of 88.6% and MCC of 0.761, both indicative of the model's good classifier ability.

Fig. 4 shows the receiver operating characteristic (ROC) curves for this final mt-QSAR linear model ML1a (Table 3). As can be observed, the model is not a random, but a truly statistically significant, classifier, since the areas under the ROC curves are significantly higher than the area under a random classifier
(= 0.5). The reliability of this final mt-QSAR linear model is further justified by the fact that the maximum inter-collinearity between two descriptors (estimated from $R^2$) found was 0.769.

In order to guarantee that the model was not developed by chance, the *Yc* randomization test was repeatedly applied 1000 times. By doing so, an average Wilk's $\lambda$ ($\lambda_{rm}$) value of 0.832 and an average accuracy value ($Acc_{rm}$) of 62.5% were obtained. Therefore, the differences between these values and those of the original ML1a model demonstrate that the latter truly constitutes a unique, not by chance, classification model for the current dataset. Following the applicability domain analysis using the standardization method [63], only three monomers, belonging to the sub-training set (*i.e.*,: monomers MD36, MD136 and MD137, see Table S1), were detected as structural outliers, although these are accurately predicted by the model.

Interpretation of any mt-QSAR model given the specific molecular features and the experimental conditions considered to the modeled toxicity is always a challenging task. The final ML1a model includes eight modified descriptors (see Eqs. (1)–(3)), three of which pertaining to the experimental element *me* while the remaining five to the experimental element *bt*. This clearly reveals that both these elements – *i.e.*,: the type of measures of toxic effects (*me*) and of biological targets (*bt*), do have an impact on the cytotoxicity of the present set of dental monomers. Further, the higher is the absolute value of a standardized LDA coefficient the higher is the weight of the independent variable in the model, therefore ranking the contributions of its descriptors as follows (see Fig. 5): $HATS5s_{bt} > R1S^{+}_{bt} >$ VE1sign_Dz(v)$_{me}$ > F10[C-O]$_{me}$ ~VE3sign_Dz(v)$_{me}$ > Eig09_EA(dm)$_{bt}$ ~$DLS\_01_{bt}$ ~$Xindex_{me}$. The meaning of these descriptors appearing in the final model is given in Table 4 along with their positive or negative contribution to the cytotoxicity endpoint response (see Table 3).

As seen in Table 4, the two most significant descriptors of the linear mt-QSAR model – *i.e.*,: HATS5s and $R1s^{+}$, are 3D-descriptors weighted by the atoms intrinsic state based on

**Table 3 – Best derived mt-QSAR classification model along with the LDA statistical results [a]**

| Model | Mt-QSAR model[b] | Sub-training set | Test set | Validation set |
|---|---|---|---|---|
| ML1a[b] | $IA_i(c_j) = +4.267 - 14.326\ VE1sign\_Dz(v)_{me} - 14.231\ F10[C-O]_{me} - 5.312\ Xindex_{me} + 26.213\ HATS5s_{bt} + 7.628\ Eig09\_EA(dm)_{bt} - 15.721\ R1s^{+}_{bt} - 9.203\ VE3sign\_B(v)_{bt} - 3.771\ DLS\_01_{bt}$<br>Wilk's $\lambda = 0.356$, $p = 6.405 \times 10^{-15}$, $F(8,79) = 17.857$, $\chi^2 = 84.67$ | TP: 50<br>TN: 33<br>FP: 2<br>FN: 3<br>*Sn*: 94.3<br>*Sp*: 94.3<br>*Acc*: 94.3<br>F1-score: 95.2<br>MCC: 0.882 | TP: 16<br>TN: 4<br>FP: 2<br>FN: 0<br>*Sn*: 66.7<br>*Sp*: 100.0<br>*Acc*: 90.9<br>F1-score: 94.1<br>MCC: 0.770 | TP: 13<br>TN: 11<br>FP: 0<br>FN: 3<br>*Sn*: 100.0<br>*Sp*: 81.3<br>*Acc*: 88.9<br>F1_score: 89.7<br>MCC: 0.799 |

[a] Metrics *Sn*, *Sp*, *Acc*, and F1-score are given in percentage. [b] Model refined from the initial ML1 model (see Fig. 1) by applying the PS3M technique.

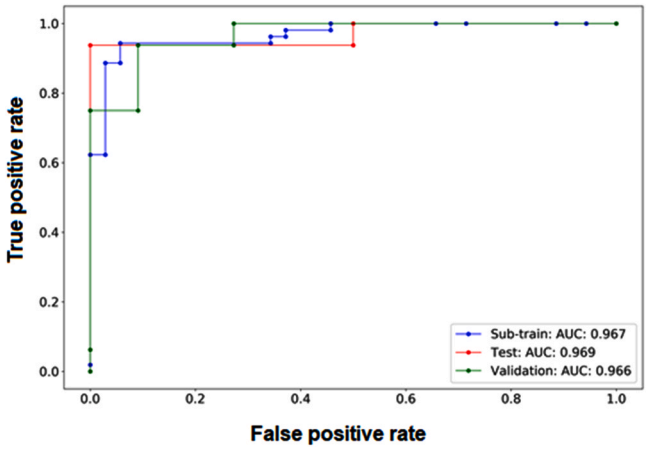


**Fig. 4 – ROC plots for the final mt-QSAR linear classification model.**

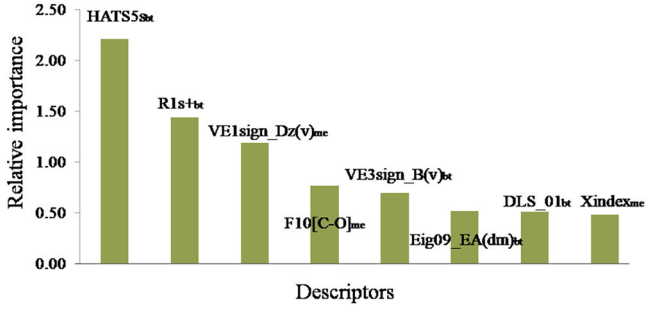


**Fig. 5 – Absolute standardized coefficients *vs.* variables in the final mt-QSAR model.**

**Table 4 – Meaning of the descriptors appearing in the final mt-QSAR linear classification model ML1a.**

| Name | Type | Description[a] | Correlation with $IA_i(c_j)$ |
|---|---|---|---|
| HATS5s | 3D-GETAWAY[b] | Leverage weighted autocorrelation of lag 5/ weighted by intrinsic state (I) | Positive |
| $R1s^+$ | 3D-GETAWAY[b] | R maximal autocorrelation of lag 1/ weighted by I-state | Negative |
| VE1sign_Dz(v) | 2D matrix-based | Coefficient sum of the last eigenvector from Barysz matrix weighted by van der Waals volume | Negative |
| F10[C-O] | 2D atom pairs | Frequency of carbon and oxygen atoms at topological distance 10 | Negative |
| VE3sign_B(v) | 2D matrix-based | Logarithm coefficient sum of the last eigenvector from Burden matrix weighted by van der Waals volume | Negative |
| Eig09_EA(dm) | Edge adjacency | Eigenvalue number 9 from edge adjacency matrix weighted by the dipole moment | Positive |
| DLS_01 | Drug-like | Modified drug-like score calculated from Lipinski's rule of five | Negative |
| Xindex | Information | Balaban X index | Negative |

[a] Refs[63,64].
[b] GETAWAY: GEometry, Topology, and Atom-Weights AssemblY.

Kier-Hall electronegativity, which is modified by the number of σ bonds, number of hydrogen atoms, number of electrons in π orbitals, and number of lone pair electrons [66,67]. Still, HATS5s and $R1s^+$ depict opposite contributions over the response variable though also depend on the kind of biological target (*bt*). The fourth most important descriptor of the model is easier to interpret – *i.e.,*: F10[C-O]. The latter descriptor accounts for the frequency of carbon and oxygen atoms located at the specific topological distance of 10, thus disclosing an important geometrical topology for the present dental monomers (acrylic acid derivatives) to turn less toxic. The remaining descriptors are either unweighted, such as the drug-like descriptor DLS_0 and the graph-based descriptor Xindex or weighted by atomic van der Waals volumes – *e.g.,*:

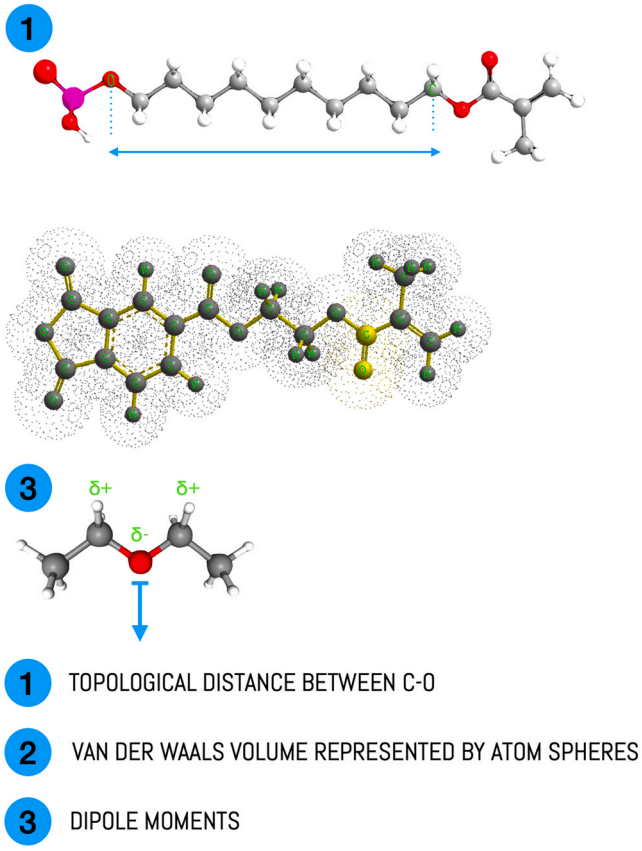


**Fig. 6 – Graphical representation of the most important descriptors in the final mt-QSAR model that were found to be responsible for the dental monomers cytotoxicity. 1. C–O topological distance; 2. van der Waals volumes; and 3. dipole moments.**

VE1sign_Dz(v) and VE3sign_B(v), and by dipole moments – *i.e.*,: Eig09_EA(dm). Concluding, one can assume that the atoms intrinsic state, van der Waals volumes, dipole moments as well as the molecular topology (see Fig. 6) could be important structural features contributing to the overall cytotoxicity of this set of dental monomers.

### 3.3. *Non-linear mt-QSAR models*

Even though the final linear mt-QSAR model demonstrated satisfactory statistical performance, it was deemed pertinent to investigate whether non-linear models may afford better predictivity. To do so, non-linear models were set up using two strategies. In the first strategy, all descriptors were employed in order to develop the model [37]. In the second one, only a limited number of descriptors obtained from feature selection was employed [43].

Using two different dataset division schemes (*i.e.*,: random and kMCA), two types of modified descriptors (Eqs. (2) and (3)), and six different machine learning techniques (*i.e.*,: kNN, NB, SVC, RF, GB and MLP), non-linear models were generated in the first stage, by employing a correlation cut-off of 0.95 and variance cut-off of 0.001 to filter out highly correlated and less-variant descriptors. Further, hyperparameter tuning was carried out for each machine learning technique to obtain the optimized parameters leading to the mostly predictive models with the sub-training set. A summary of the statistical performance of these non-linear classification models (MNL1-MNL24) is provided in Table S6 of the Supplementary Material. Then, similarly, non-linear models were developed using rather the eight descriptors included in the most predictive linear model ML1a as well as its data-distribution. A summary of the statistical performance of the latter non-linear classification models (MNL25-MNL30) is provided in Table S7 of the Supplementary Material. As can be judged, both the results in Tables S6 and S7 indicate that the MLP technique stands up as the best predictor classifier for the cytotoxicity response. Yet as far as overall predictivity is concerned, the MLP model based on a limited number of descriptors (*i.e.*, MNL27 of Table S7) is found to be more predictive than the MLP model based on a maximum descriptor space (*i.e.*, MNL9 of Table S4). Nevertheless, since both these MLP models were generated with a fixed hidden layer size of (100,100), one needs to further check if more predictive models could be produced by varying the hidden layer size. Starting from the optimized parameters of MNL9 and MNL27, additional MLP models were thus set up with different hidden layer sizes, the results of which are summarized in Table S8 of the Supplementary Material. As can be observed, model MNL27b generated with a hidden layer size of (50,0)

**Table 5 – Results of condition-wise prediction using the best mt-QSAR linear classification model (model ML1a; Table 3**

| Set | Condition | *me* | *bt* | #Instances | Accuracy (%) |
|---|---|---|---|---|---|
| Sub-training and test sets | CX01 | $EC_{50}$ | A549 | 1 | 100.0 |
| | CX02 | $ED_{50}$ | 3T3 | 7 | 71.4 |
| | CX03 | $ED_{50}$ | HGF | 11 | 100.0 |
| | CX04 | $ED_{50}$ | HPF | 8 | 100.0 |
| | CX05 | $ED_{50}$ | HPLF | 7 | 100.0 |
| | CX06 | $IC_{50}$ | A549 | 1 | 100.0 |
| | CX07 | $IC_{50}$ | HGF | 1 | 100.0 |
| | CX08 | $IC_{50}$ | Hela S3 | 32 | 93.8 |
| | CX09 | $IC_{50}$ | JTC-12 | 8 | 87.5 |
| | CX10 | $IC_{50}$ | PBM | 2 | 100.0 |
| | CX11 | $LC_{50}$ | RAW264.7 | 4 | 75.0 |
| | CX12 | $LC_{50}$ | Rat Hepatocytes | 8 | 100.0 |
| | CX13 | $TC_{50}$ | BEAS-2B | 1 | 100.0 |
| | CX14 | $TC_{50}$ | C3H10T1/2 | 1 | 100.0 |
| | CX15 | $TC_{50}$ | HGF | 1 | 100.0 |
| | CX16 | $TC_{50}$ | L929 | 6 | 100.0 |
| | CX17 | $TC_{50}$ | MC3T3-E1 | 1 | 100.0 |
| | CX18 | $TC_{50}$ | PBM | 3 | 100.0 |
| | CX19 | $TC_{50}$ | RPC-C2A | 3 | 100.0 |
| | CX20 | $TC_{50}$ | THP-1 | 2 | 50.0 |
| | CX21 | $TC_{50}$ | V79–4 | 2 | 100.0 |
| Validation set | CX01 | $EC_{50}$ | A549 | 2 | 100.0 |
| | CX02 | $ED_{50}$ | 3T3 | 3 | 100.0 |
| | CX03 | $ED_{50}$ | HGF | 2 | 100.0 |
| | CX04 | $ED_{50}$ | HPF | 2 | 100.0 |
| | CX05 | $ED_{50}$ | HPLF | 3 | 66.7 |
| | CX08 | $IC_{50}$ | Hela S3 | 7 | 100.0 |
| | CX09 | $IC_{50}$ | JTC-12 | 1 | 0.0 |
| | CX10 | $IC_{50}$ | PBM | 1 | 100.0 |
| | CX12 | $LC_{50}$ | Rat Hepatocytes | 2 | 100.0 |
| | CX16 | $TC_{50}$ | L929 | 2 | 50.0 |
| | CX20 | $TC_{50}$ | THP-1 | 1 | 100.0 |
| | CX21 | $TC_{50}$ | V79–4 | 1 | 100.0 |

should be considered as the best model, considering its external predictivity. However, the maximum overall predictivity was undoubtedly obtained with MNL27f, generated with a hidden layer size of (50,50). Even so, none of these non-linear MLP models was able to match the overall predictivity of the previous linear model ML1a (see Table S8). Therefore, one can assume that the refined model ML1a is not only the best linear model, but also the most predictive mt-QSAR classification model.

Finally, a 'condition-wise prediction' analysis of the mt-QSAR model ML1a was carried out. This analysis simply tests the results of the model and breaks them on the basis of the accuracy obtained against the different conditions. The results of such analysis are presented in Table 5.

As can be judged, model ML1a clearly shows consistently high accuracy values against most of the targeted experimental conditions. This means that the model is capable of predicting the cytotoxicity irrespectively of both the kind of cell lines and different type of measures of such endpoint response. One should notice nevertheless that for some conditions such as CX20 (*me*: $TC_{50}$, *bt*: THP-1), CX09 (*me*: $IC_{50}$, *bt*: JTC-12) and CX16 (*me*: $IC_{50}$, *bt*: L929), less than 60% accuracy was obtained for either the modeling set (sub-training plus test sets) or the validation set.

### 3.4. *Additional challenges and future applications*

The main goal of this work was to develop both single and multi-target models that could predict and characterize the cytotoxic behavior of acrylic acid-based dental monomers following the OECD guidelines. As depicted in Fig. 1, the st-QSAR model could successfully explain the similarity relationships between active monomers and their biological potency differences. However, the smaller number of data-points and their lesser structurally diversity may restrict its overall applicability. In contrast, the applied mt-QSAR modeling is based on a classification analysis and *per se* it does not allow to predict the biological potency values. Yet, such mt-QSAR models are extremely useful since they are able to describe biological potency differences between active and inactive compounds with respect to their structures. Besides, the present mt-QSAR classification model offers several advantages. For example, by incorporating multiple experimental assay conditions (as shown in Table 5), one is able to explain their influence in the biological potency differences among active and inactive data-points. It also helped in improving the applicability domain of the models to a considerable extent by including a larger number of structurally diverse compounds. More importantly such mt-approach

affords a unique model that may be used for predicting the activity under any experimental condition present in the sub-training set.

The models set up in the present work will serve as important guidelines for designing or selecting novel acrylic acid-based dental monomers. In dental materials research, the workflow that is followed during the design of direct restorative polymers could be substantially modernized by resorting to these models. In fact, when inclusion of new monomers is planned for experimental dental polymer formulations, biocompatibility studies are one of the first steps taken. Due to their importance, cytotoxicity studies decide whether it is feasible to advance, or not, with the new formulations [68]. Therefore, models such as the ones presented here may significantly advance and improve the workflow of dental materials design, reducing the time needed while minimizing resources. Also, it is expected that additional experimental cytotoxicity data of other dental monomers will be undertaken, expanded, and reported in future scientific literature. Thus, the current models may be further refined based on these additional data.

The present work also confirmed that multi-target QSAR modeling based on the Box-Jenkins moving average approach may work well even with a comparatively smaller number of datasets, judging from the high predictive accuracy of the mt-QSAR models as well as from the results of the condition-wise prediction scheme. However, more research with a larger number of datasets is required to understand the true significance of the present multi-target QSAR modeling in handling this type of problems. In addition, our newly proposed PS3M analysis was found to be a very effective optimization strategy for improving the quality of linear regression and classification models. Further research in this area will most certainly help advancing QSAR model development for big datasets.

## 4. Conclusions

In this work, it was possible to successfully set up predictive QSAR models able to disclose the structural attributes of acrylic acid based dental monomers responsible for higher cytotoxicity towards multiple cell lines. Initially, a single-target model was built with 39 compounds from cytotoxicity MTT assays reported against Hela S3 cell line. Following, by merging such single-target data with additional 99 compounds assayed against 17 other different cell lines, an attempt was made to build multi-target QSAR classification models targeting such heterogeneous dataset. The best-fit multi-target QSAR linear model was found to be the most predictive one, even when compared to non-linear models developed with various machine learning tools. This linear mt-QSAR model, with a particularly good predictive accuracy for discriminating among dental monomers (> 90%), revealed also the crucial structural features responsible for their higher cytotoxicity. As expected, the single target model though simpler than the multi-target model depicted the importance of some atom-centered fragments for the cytotoxicity against Hela S3 cell lines. Even if no common descriptors were found in these single and multi-target models, some common inferences are worthy pointing out. For example, according to both single and multi-target models, atomic van der Waals volume could play a key role in inducing the cytotoxicity. Additionally, the multi-target models revealed the relevance as well for their cytotoxicity of other features such as the topological distance between carbon and oxygen atoms, dipole moments, plus that of the atom's intrinsic states. What is more, the importance of considering both experimental elements (*i.e.*, type of measure and biological target) to successfully target the monomers' cytotoxicity was clearly denoted. All the models were set up using non-commercial and open-access platforms, and therefore these are easily reproducible. Apart from describing the mt-QSAR modeling of the acrylic acid based dental monomers for the first time, this work is reporting a new technique of analysis and *in house* tool named PS3M, which was placed in the public domain. This technique was found to be very useful for ascertaining the predictive quality of both regression and classification models. It is thus anticipated that the technique and the tool shall play a significant role in advancing linear mt-QSAR model development in the future.

To conclude, the overall information then gathered and the QSAR models *per se* can be used as reliable screening tools, valuable for future planning and design of novel bulk monomers or functional monomers to be included in resin composites, dental adhesives, or other polymer formulations used in dentistry.

## Acknowledgment

This work received financial support from Fundação para a Ciência e a Tecnologia (FCT/MCTES) through national funds (LAQV-REQUIMTE, grant UID/QUI/50006/2020).

## Appendix A. Supporting information

Supplementary data associated with this article can be found in the online version at doi:10.1016/j.dental.2021.12.014.

REFERENCES

[1] Barszczewska-Rybarek IM. A Guide through the dental dimethacrylate polymer network structural characterization and interpretation of physico-mechanical properties. Materials 2019;12. http//doi.org/ARTN405710.3390/ma12244057.
[2] Moszner N, Hirt T. New polymer-chemical developments in clinical dental polymer materials: enamel-dentin adhesives and restorative composites. J Polym Sci Pol Chem 2012;50:4369–402. https://doi.org/10.1002/pola.26260
[3] Hikage S, Sato A, Suzuki S, Cox CF, Sakaguchi K. Cytotoxicity of dental resin monomers in the presence of S9 mix enzymes. Dent Mater J 1999;18:76–86. https://doi.org/10.4012/dmj.18.76
[4] Moharamzadeh K, Brook IM, Van Noort R. Biocompatibility of resin-based dental materials. Materials 2009;2:514–48. https://doi.org/10.3390/ma2020514
[5] Kim EC, Park H, Lee SI, Kim SY. Effect of the acidic dental resin monomer 10-methacryloyloxydecyl dihydrogen phosphate on odontoblastic differentiation of human dental pulp cells. Basic Clin Pharm 2015;117:340–9. https://doi.org/10.1111/bcpt.12404

[6] Goldberg M. In vitro and in vivo studies on the toxicity of dental resin components: a review. Clin Oral Investig 2008;12:1–8. https://doi.org/10.1007/s00784-007-0162-8

[7] Jun SK, Cha JR, Knowles JC, Kim HW, Lee JH, Lee HH. Development of Bis-GMA-free biopolymer to avoid estrogenicity. Dent Mater 2020;36:157–66. https://doi.org/10.1016/j.dental.2019.11.016

[8] Laurent P, Camps J, De Meo M, Dejou J, Abouta I. Induction of specific cell responses to a Ca3SiO5-based posterior restorative material. Dent Mater 2008;24:1486–94. https://doi.org/10.1016/j.dental.2008.02.020

[9] Yourtee D, Holder AJ, Smith R, Morrill JA, Kostoryz E, Brockmann W, et al. Quantum mechanical quantitative structure activity relationships to avoid mutagenicity in dental monomers. J Biomat Sci-Polym E 2001;12:89–105. https://doi.org/10.1163/156856201744470

[10] Reichl FX, Simon S, Esters M, Seiss M, Kehe K, Kleinsasser N, et al. Cytotoxicity of dental composite (co)monomers and the amalgam component Hg2+ in human gingival fibroblasts. Arch Toxicol 2006;80:465–72. https://doi.org/10.1007/s00204-006-0073-5

[11] Leprince JG, Palin WM, Hadis MA, Devaux J, Leloup G. Progress in dimethacrylate-based dental composite technology and curing efficiency. Dent Mater 2013;29:139–56. https://doi.org/10.1016/j.dental.2012.11.005

[12] Gupta SK, Saxena P, Pant VA, Pant AB. Release and toxicity of dental resin composite. Toxicol Int 2012;19:225–34. https://doi.org/10.4103/0971-6580.103652

[13] Neves SO, Magalhaes LMD, Correa JD, Dutra WO, Gollob KJ, Silva TA, et al. Composite-derived monomers affect cell viability and cytokine expression in human leukocytes stimulated with Porphyromonas gingivalis. J Appl Oral Sci 2019;27:e20180529. https://doi.org/ARTN e2018052910.1590/1678-7757-2018-0529.

[14] Hashimoto M, Ohno H, Sano H, Kaga M, Oguchi H. In vitro degradation of resin-dentin bonds analyzed by microtensile bond test, scanning and transmission electron microscopy. Biomaterials 2003;24:3795–803. https://doi.org/10.1016/S0142-9612(03)00262-X

[15] Issa Y, Watts DC, Brunton PA, Waters CM, Duxbury AJ. Resin composite monomers alter MTT and LDH activity of human gingival fibroblasts in vitro. Dent Mater 2004;20:12–20. https://doi.org/10.1016/S0109-5641(03)00053-8

[16] Bandarra S, Mascarenhas P, Luis AR, Catrau M, Bekman E, Ribeiro AC, et al. In vitro and in silico evaluations of resin-based dental restorative material toxicity. Clin Oral Investig 2020;24:2691–700. https://doi.org/10.1007/s00784-019-03131-4

[17] Bakopoulou A, Papadopoulos T, Garefis P. Molecular toxicology of substances released from resin-based dental restorative materials. Int J Mol Sci 2009;10:3861–99. https://doi.org/10.3390/ijms10093861

[18] Murray PE, Garcia Godoy C, Garcia Godoy F. How is the biocompatibilty of dental biomaterials evaluated? Med Oral Patol Oral Cir Bucal 2007;12:E258–66.

[19] Pedano MS, Li X, Yoshihara K, Van Landuyt K, Van Meerbeek B. Cytotoxicity and bioactivity of dental pulp-capping agents towards human tooth-pulp cells: a systematic review of in-vitro studies and meta-analysis of randomized and controlled clinical trials. Materials 2020;13. https://doi.org/10.3390/ma13122670. (2670).

[20] Yoshii E. Cytotoxic effects of acrylates and methacrylates: relationships of monomer structures and cytotoxicity. J Biomed Mater Res 1997;37:517–24. https://doi.org/10.1002/(Sici)1097-4636(19971215)37:4<517::Aid-Jbm10>3.0.Co;2-5

[21] Thonemann B, Schmalz G, Hiller KA, Schweikl H. Responses of L929 mouse fibroblasts, primary and immortalized bovine dental papilla-derived cell lines to dental resin components. Dent Mater 2002;18:318–23. https://doi.org/10.1016/S0109-5641(01)00056-2

[22] Caldas IP, Alves GG, Barbosa IB, Scelza P, de Noronha F, Scelza MZ. In vitro cytotoxicity of dental adhesives: a systematic review. Dent Mater 2019;35:195–205. https://doi.org/10.1016/j.dental.2018.11.028

[23] Manganelli S, Leone C, Toropov AA, Toropova AP, Benfenati E. QSAR model for predicting cell viability of human embryonic kidney cells exposed to SiO2 nanoparticles. Chemosphere 2016;144:995–1001. https://doi.org/10.1016/j.chemosphere.2015.09.086

[24] Muratov EN, Bajorath J, Sheridan RP, Tetko IV, Filimonov D, Poroikov V, et al. QSAR without borders. Chem Soc Rev 2020;49:3525–64. https://doi.org/10.1039/d0cs90041a

[25] Holder AJ, Ye L, Yourtee DM, Agarwal A, Eick JD, Chappelow CC. An application of the QM-QSAR method to predict and rationalize lipophilicity of simple monomers. Dent Mater 2005;21:591–8. https://doi.org/10.1016/j.dental.2004.08.004

[26] Perez-Garrido A, Helguera AM, Rodriguez FG, Cordeiro MNDS. QSAR models to predict mutagenicity of acrylates, methacrylates and alpha,beta-unsaturated carbonyl compounds. Dent Mater 2010;26:397–415. https://doi.org/10.1016/j.dental.2009.11.158

[27] Neves BJ, Braga RC, Melo-Filho CC, Moreira-Filho JT, Muratov EN, Andrade CH. QSAR-based virtual screening: advances and applications in drug discovery. Front Pharm 2018;0:1275. https://doi.org/10.3389/FPHAR.2018.01275

[28] Burello E. Review of (Q)SAR models for regulatory assessment of nanomaterials risks. NanoImpact 2017;8:48–58. https://doi.org/10.1016/J.IMPACT.2017.07.002

[29] Cumming JG, Davis AM, Muresan S, Haeberlein M, Chen H. Chemical predictive modelling to improve compound quality. Nat Rev Drug Discov 2013;12:948–62. https://doi.org/10.1038/nrd4128

[30] Speck-Planche A, Cordeiro MNDS. Advanced in silico approaches for drug discovery: mining information from multiple biological and chemical data through mtkQSBER and pt-QSPR strategies. Curr Med Chem 2017;24:1687–704. https://doi.org/10.2174/0929867324666170124152746

[31] Speck-Planche A. Combining ensemble learning with a fragment-based topological approach to generate new molecular diversity in drug discovery: in silico design of Hsp90 inhibitors. ACS Omega 2018;3:14704–16. https://doi.org/10.1021/acsomega.8b02419

[32] Halder AK, Cordeiro MNDS. Development of multi-target chemometric models for the inhibition of class I PI3K enzyme isoforms: a case study using QSAR-Co tool. Int J Mol Sci 2019:20. https://doi.org/10.3390/ijms20174191

[33] Ambure P, Halder AK, Diaz HG, Cordeiro MNDS. QSAR-Co: an open source software for developing robust multitasking or multitarget classification-based qsar models. J Chem Inf Model 2019;59:2538–44. https://doi.org/10.1021/acs.jcim.9b00295

[34] Sushko I, Novotarskyi S, Korner R, Pandey AK, Rupp M, Teetz W, et al. Online chemical modeling environment (OCHEM): web platform for data storage, model development and publishing of chemical information. J Comput Aided Mol Des 2011;25:533–54. https://doi.org/10.1007/s10822-011-9440-2

[35] Sadowski J, Gasteiger J, Klebe G. Comparison of automatic 3-dimensional model builders using 639 x-ray structures. J Chem Inf Comp Sci 1994;34:1000–8. https://doi.org/10.1021/ci00020a039

[36] Raschka S. MLxtend: providing machine learning and data science utilities and extensions to Python's scientific computing stack. J Open Source Softw 2018;3. (638).

[37] Halder AK. Finding the structural requirements of diverse HIV-1 protease inhibitors using multiple QSAR modelling for lead identification. SAR QSAR Environ Res 2018;29:911–33. https://doi.org/10.1080/1062936X.2018.1529702

[38] Tetko IV, Tanchuk VY, Villa AE. Prediction of n-octanol/water partition coefficients from PHYSPROP database using artificial neural networks and E-state indices. J Chem Inf Comput Sci 2001;41:1407–21. https://doi.org/10.1021/ci010368v
[39] Roy PP, Paul S, Mitra I, Roy K. On two novel parameters for validation of predictive QSAR models. Molecules 2009;14:1660–701. https://doi.org/10.3390/molecules14051660
[40] Golbraikh A, Tropsha A. Beware of q2!. J Mol Graph Model 2002;20:269–76. https://doi.org/10.1016/s1093-3263(01)00123-1
[41] Halder AK, Giri AK, Cordeiro MNDS. Multi-target chemometric modelling, fragment analysis and virtual screening with ERK inhibitors as potential anticancer agents. Molecules 2019;24. https://doi.org/10.3390/molecules24213909. (3909).
[42] Halder AK, Cordeiro MNDS. Probing the environmental toxicity of deep eutectic solvents and their components: an in silico modeling approach. ACS Sustain Chem Eng 2019;7:10649–60. https://doi.org/10.1021/acssuschemeng.9b01306
[43] Speck-Planche A. Multicellular target QSAR Model for simultaneous prediction and design of anti-pancreatic cancer agents. ACS Omega 2019;4:3122–32. https://doi.org/10.1021/acsomega.8b03693
[44] Speck-Planche A. Multi-scale QSAR approach for simultaneous modeling of ecotoxic effects of pesticides. In: Roy K, editor. Ecotoxicological QSARs. New York, NY: Springer US; 2020. p. 639–60.
[45] Speck-Planche A, Scotti MT. BET bromodomain inhibitors: fragment-based in silico design using multi-target QSAR models. Mol Divers 2019;23:555–72. https://doi.org/10.1007/s11030-018-9890-8
[46] Halder AK, Cordeiro MNDS. QSAR-Co-X: an open source toolkit for multi-target QSAR modelling. J Chemin-2021;13:29. https://doi.org/10.1186/s13321-021-00508-0
[47] Gore PA. Chapter 11 - Cluster analysis. In: Tinsley HEA, Brown SD, editors. Handbook of Applied Multivariate Statistics and Mathematical Modeling. San Diego: Academic Press; 2000. p. 297–321.
[48] Brown MT, Wicker LR. Chapter 8 - Discriminant analysis. In: Tinsley HEA, Brown SD, editors. Handbook of Applied Multivariate Statistics and Mathematical Modeling. San Diego: Academic Press; 2000. p. 209–35.
[49] Wilks SS. Certain generalizations in the analysis of variance. Biometrika 1932;24:471–94.
[50] Boughorbel S, Jarray F, El-Anbari M. Optimal classifier for imbalanced data using Matthews correlation coefficient metric. PLOS One 2017:12. https://doi.org/10.1371/journal.pone.0177678
[51] Fawcett T. An introduction to ROC analysis. Pattern Recognit Lett 2006;27:861–74. https://doi.org/10.1016/j.patrec.2005.10.010
[52] Hanczar B, Hua JP, Sima C, Weinstein J, Bittner M, Dougherty ER. Small-sample precision of ROC-related estimates. Bioinformatics 2010;26:822–30. https://doi.org/10.1093/bioinformatics/btq037
[53] Douglass MJJ. Hands-on machine learning with Scikit-Learn, Keras, and Tensorflow, 2nd edition. Phys Eng Sci Med 2020;43:1135–6. https://doi.org/10.1007/s13246-020-00913-z
[54] Cover TM, Hart PE. Nearest neighbor pattern classification. IEEE Trans Inf Theory 1967;13:21–7. https://doi.org/10.1109/Tit.1967.1053964
[55] McCallum A, Nigam K. A comparison of event models for naive bayes text classification. Work Learn Text Categ 2001;752.
[56] Boser BE, Guyon, I.M., & Vapnik, V.N. A training algorithm for optimal margin classifiers. In Proceedings of the Fifth Annual Workshop on Computational Learning Theory ACM, pp. 144–152.
[57] Breiman L. Random forests. Mach Learn 2001;45:5–32. https://doi.org/10.1023/A:1010933404324
[58] Friedman JH. Greedy function approximation: a gradient boosting machine. Ann Stat 2001;29:1189–232. https://doi.org/10.1214/aos/1013203451
[59] Huang GB, Babri HA. Upper bounds on the number of hidden neurons in feed forward networks with arbitrary bounded nonlinear activation functions. IEEE Trans Neural Netw Learn Syst 1998;9:224–9. https://doi.org/10.1109/72.655045
[60] Ojha PK, Roy K. Comparative QSARs for antimalarial endochins: importance of descriptor-thinning and noise reduction prior to feature selection. Chemom Intell Lab 2011;109:146–61. https://doi.org/10.1016/j.chemolab.2011.08.007
[61] Gramatica P. On the development and validation of QSAR models. Methods Mol Biol 2013;930:499–526. https://doi.org/10.1007/978-1-62703-059-5_21
[62] Serra A, Onlu S, Festa P, Fortino V, Greco D. MaNGA: a novel multi-niche multi-objective genetic algorithm for QSAR modelling. Bioinformatics 2020;36:145–53. https://doi.org/10.1093/bioinformatics/btz521
[63] Roy K, Kar S, Ambure P. On a simple approach for determining applicability domain of QSAR models. Chemom Intell Lab 2015;145:22–9. https://doi.org/10.1016/j.chemolab.2015.04.013
[64] Khan PM, Roy K. Current approaches for choosing feature selection and learning algorithms in quantitative structure-activity relationships (QSAR). Expert Opin Drug Dis 2018;13:1075–89. https://doi.org/10.1080/17460441.2018.1542428
[65] Ambure P, Aher RB, Gajewicz A, Puzyn T, Roy K. "NanoBRIDGES" software: open access tools to perform QSAR and nano-QSAR modeling. Chemom Intell Lab 2015;147:1–13. https://doi.org/10.1016/j.chemolab.2015.07.007
[66] Todeschini R, Consonni V, Todeschini R. Molecular Descriptors for Chemoinformatics. second ed. Weinheim Chichester: Wiley-VCH; John Wiley distributor,; 2009.
[67] Todeschini R, Consonni V. Handbook of Molecular Descriptors. Weinheim; New York: Wiley-VCH,; 2000.
[68] Shen C, Rawls RH, Esquivel-Upshaw JF. In vitro research on dental materials. In: Shen C, Rawls RH, Esquivel-Upshaw JF, editors. Phillip's Science of Dental Materials. New York: Saunders; 2021. p. 381–9.