Computer-Aided Classification of New Psychoactive Substances

Alina Bărbulescu,1 Lucica Barbeș,2 and Cristian-Ștefan Dumitriu3

1Transylvania University of Brașov, Department of Civil Engineering, 5 Turnului Str., Brașov, Romania
2Ovidius University of Constanta, Department of Chemistry and Chemical Engineering, 124 Mamaia Bd., Constanta, Romania
3SC Utilnavorep SA, 55 Aurel Vlaicu Av., Constanța 90055, Romania

Correspondence should be addressed to Lucica Barbeș; lucille.barbes2009@gmail.com

Received 17 November 2021; Accepted 9 December 2021; Published 31 December 2021

Academic Editor: Hassan Arida

Copyright © 2021 Alina Bărbulescu et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

The appearance on the free market of synthetic cannabinoids raised the researchers’ interest in establishing their molecular similarity by QSAR analysis. A rigorous criterion for classifying drugs is their chemical structure. Therefore, this article presents the structural similarity of two groups of drugs: benzoylindoles and phenylacetylindoles. Statistical analysis and clustering of the molecules are performed based on their numerical characteristics extracted using Cheminformatics methods. Their similarities/dissimilarities are emphasized using the dendrograms and heat map. The highest discrepancies are found in the phenylacetylindoles group.

1. Introduction

The consumption of psychoactive substances is among the principal causes of young people’s mortality and health issues. A significant proportion of adults consumed some drugs at a certain period in their lives or presented with drug addiction. Synthetic cannabinoids are found in “herbal smoking mixes” and mixtures with other psychoactive substances such as sedatives/hypnotics, hallucinogens, and stimulants [1]. Manufactured in clandestine laboratories and widely sold, the class of new psychoactive substances known as different spices poses a growing threat to public health, such as Spice, Spice Gold, and Silver. In some countries, acute poisonings of users’ groups [2–4] have been recorded. Synthetic cannabinoids have been missold as THC (∆9-tetrahydrocannabinol) or CBD (cannabidiol) and are packed in tablets containing powder or utilized in devices as liquids [5].

Another recent development is the popularity of “vaping” among young people, synthetic cannabinoids being discovered in cartridges filled with liquid for use in e-cigarettes [6].

Psychoactive drugs increase the dopamine available in the brain (called the reward circuit) with participation in pleasure modulation. Synthetic cannabinoids (SCs) are much stronger and more toxic than users can often think. Their increasing prevalence has led to severe health problems among young people and adults. Consumption of these new psychoactive substances poses a health risk and generates severe issues in adapting to reality, education, work, family life, incidents, and road accidents [7, 8]. Research has been performed on the cannabinoids detected in different products to determine how cannabinoids affect the human body [9]. They also aim to analyze the possibilities of using these substances to treat certain neurodegenerative diseases, drug addiction, emotional disorders, or cancer [10]. However, separating the psychoactive effects and the medicinal properties of these drugs is challenging [11].

In about 8–10 years, researchers have tried different dynamic simulation models on rapid identification and establishing the relationships of addiction between the chemical structure and effects on consumers to detect possible ways for curing addiction and overdose [12]. Functionally similar to THC (the cannabis active principle), SCs bind to the same receptors in the human body [6]. The group of synthetic cannabinoids or synthetic cannabinimimetic is the largest monitored by the EU Early Warning System. In December 2016, 169 synthetic cannabinoids were reported to the EMCDDA and increased to 190
substances reported in 2018 [8]. About 280 synthetic cannabinoids were known at the end of 2019 [6].

SCs fall into different structural groups [13, 14], such as aminoalkylindoles (which include naphthoylindoles, phenylacetylamides, naphthylmethylindoles, and benzoylindoles), classical cannabinoids, nonclassical cannabinoids, hybrid cannabinoids, eicosanoids, benzimidazole, indazole-based SCs, and others (naphthoylepyrroles, naphthylmethylindenes, carboxamide-type synthetic cannabinoids, etc.) [15, 16]. The first three groups are presented in Figure 1.

In the investigation of the various chemical synthesis compounds or any chemical synthesis preparation, the chemical analyst establishes the relationships between physicochemical parameters and biological (pharmacological) properties, the degree of structural superposability of the chemical products, using filters based on computational structure and quantitative relationship models of structure-activity (QSARs models) [11].

The search for similarity using molecular fingerprints are among the most common approaches in identifying active chemical compounds proposed for analysis in this study. Many structural descriptors or 2D and 3D ownership are encoded in different representations of fingerprints. All fingerprints eventually transform molecular likeness analysis into a paired comparison of template compound patterns and the entire database [17].

Several fingerprint types are known, each representing a difference in the molecule. Artificial intelligence techniques have been employed in the last period for data analysis in various domains [18–21]. Cheminformatics is a tool used to analyze statistical data based on the digital footprint, which is the dynamic field of modern design for pharmaceuticals, in particular. Cheminformatics plays an essential role in accumulating, grouping, examining chemical data, and finding new entities based on which new chemical structures can build active molecules [22].

Swandana et al. [23] showed that the pharmacokinetic parameters could be predicted using the in silico method. The authors emphasize some software (admetSAR, Cheminform, Molinspiration, QikProp, SwissADME, and pkCSM software) to predict oral-systemic drugs’ pharmacokinetics characteristics. Other scientists [24, 25] emphasized different computational methods utilized for discovering drugs and their limitations.

Finding similarities of different molecules has been of interest since the 90s [26] that can be easily done using the new capabilities of the new software. The rcdk, ChemmineR, and rpubchem packages of R are powerful tools in the field of cheminformatics [27–32]. Their utilization helps the cheminformatics-oriented R user to organize chemical information efficiently. Guha et al. [27] showed that one of the main tasks is to include further coverage of the CDK API by developing the CDK itself, such as extending the 3D builder model. The authors noted that the CDK API could be accessed directly via rJava, so the functions provided by the rcdk package make cheminformatics functionality readily available for the end-users.

In this study, we used the R software capacity to characterize 14 new cannabinoids from the benzoylindoles and phenylacetylamides groups and determine the similarities among their molecules. Clustering techniques are applied for grouping the studied molecules based on the determined characteristics to cross-validate similar attributes in the same group of drugs.

2. Materials and Methods

The structures subject to the study have been downloaded as .sdf files from PubChem [33]. They belong to the class aminoalkylindoles, subclasses benzoylindoles, and phenylacetylamides.

The R software has been used for performing the study. For this aim, the following libraries have been installed: ChemmineR, chemometrics, cluster, factoextra, fingerprint, fncsR, ggplot2, gridExtra, iqspr, NbClust, rcdk, rJava, rgl, and vegan. They are necessary for drawing the structures, finding the characteristics of the chemical structures, or determining the appurtenance of the chemical structures to some clusters.

The stages of this study are presented in Figure 2. They are

(i) Import the molecule from ChemPub, as SDF files, and visualize them
(ii) Visualize the molecule structure and molecular formula (MF), and determine the molar weight filter (MW), the types and numbers of atoms, and the functional groups together with the corresponding frequencies
(iii) Compute different descriptors utilizing the library ChemmineOB:
   (i) HBA1: number of hydrogen bond acceptors 1
   (ii) HBA2: number of hydrogen bond acceptors 2
   (iii) HBD: number of hydrogen bond donors
   (iv) logP: logarithmic of partition coefficients (log P)
   (v) MR: molar refractivity
   (vi) TPSA: topological polar surface area

The hydrogen bonds are among the most significant interactions between the solute and solvent. In binding the ligands, the hydrogen bonds have important contributions to the substrates, antagonists, and agonist recognition [34]. The octanol/water partition coefficient is used to characterize the lipophilicity of the compounds and is defined by the ratio \( \text{[solute]}_{\text{octanol}} / \text{[solute]}_{\text{water}} \). Generally, the logarithm of this ratio, \( \log P \), is reported [35].

The molar refractivity (MR) measures the steric factor, which is the volume covered by an atom or a group of atoms. It is correlated to the London dispersive forces acting during the drug-receptor interaction [36].

The hydrophilic interaction and the hydrogen bond formation are reflected by the polar surface (PSA) [37], computed by eliminating the area of nonpolar hydrogen, carbon, and halogen atoms from the molecular surface. Determined by the method of Ertl et al. [38], it is referred to
According to the available literature, Turner and Agatonovic-Kustrin [40] showed that most drugs have a molar weight lower than 500 and PSA less than 120 Å.

(i) Search the atom pair similarity by cmp.similarity command from ChemmineR. For this aim, the Tanimoto index was utilized [41] due to its performance [26]. For dichotomous variables, it is computed by

$$SS_{A,B} = \frac{c}{(a + b - c)}$$

The following relationship exists between the distance and similarity metrics:

$$similarity = \frac{1}{(1 + distance)}$$

(ii) Determine the intermolecular distances, using the fmcs function from the fmiscR library and the compound similarities.

(iii) Clustering the molecules’ sets: for this aim, the binning clustering (using cmp.function that implements the Tanimoto and Tversky methods) [41, 42], the Jarvis–Patrick algorithm [43], the k-means, and the hierarchical clustering (with the NbClust package) have been utilized [44]. Comparisons are provided.

3. Results and Discussion

The structures of the molecules from the benzoylindoles group (Group 1) are represented in Figure 3, and those from the phenylacetylindoles (Group 2) are represented in Figure 4, together with the identification number from PubChem.

A similar representation, but showing the number of atoms, is presented in Figure 5, for the molecule 9889172. It contains an automatic numbering added by R software when extracting the information from PubChem: the atoms are numbered consecutively (from 1 to the total number of atoms), based on the following order: elements in the seventh group, from the highest to the lowest atomic number, elements in the sixth group, from the highest to the lowest atomic number, and so on, from left to right in the Mendeleev table.

Table 1 contains the molecules formula, the types and numbers of atoms in each molecule from both groups, the molecular weight filters (MW), the number of rings, and the number of aromatic rings retrieved using ChemmineR. The molar weight filters vary in a larger range in the first group by comparison to the second one.

All but the first molecule do not contain F; all the molecules in Group 2 do not contain I. Only one molecule in each group contains Cl. The number of N and O is one or two. The number of C and H atoms is lower in the first group than the second one.

The analysis shows that no molecule contains groups RH2, R2NH, ROPO3, ROH, RCHO, RCOOH, RCOOR, RCCH, and RCN, but all contain rings, most of them aromatic.

Table 2 contains the most important descriptors that characterize the molecules. The values of HBA1 are lower for the first group, HBD is absent for both groups, logP is in the range of 3.4594–5.8860 for Group 1 and 4.4975–6.0457 for Group 2 (showing higher hydrophilicity for the second group compared to the first one). TPSA varies in larger limits for Group 1 than for Group 2. The higher the TPSA is, the higher the drug transport is. In the studied case, the molecule with CID = 56463 has the highest TPSA.
The similarity among the pairs of atoms in different molecules is given in Table 3.

Most similarity coefficients are above 0.5, the best ones corresponding to the pair of CIDs (57507911, 57507905), (988917, 0.8462) in Group 1 and (44397540, 11616723), (44397500, 11616723) in Group 2.

The compounds' similarity has been studied using the fmcs function. The output shows the size of the query molecule and that of the target molecule (Query_Size and Target_Size, respectively), the Tanimoto coefficient, and the overlap coefficient, measuring the structures' superposition degree.

For example, in Table 4, the query molecule has CID 9889172, and the target molecule in row three has CID 117587582. For both, the query and target sizes are equal (24 atoms, the number of hydrogen atoms is not counted), the Tanimoto coefficient is 0.92, and the overlap coefficient is 0.9583. The higher the Tanimoto and the overlap coefficients are, the higher the similarity of the structures is. The highest dissimilarity is that between molecule CID 56463 and the other molecules in Group 1.

The similarity coefficients in Group 2 are higher than in the first group, as shown in Table 5. Hence, the compounds in the second group have more similar structural characteristics than those in the first group.

The computed similarity coefficients between the pairs of atoms in different molecules from Group 1 and those from Group 2 are shown in Table 6.

The similarities of atom pairs from 9889172 (Group 1) with those from 44397500 and 1161723 (Group 2) are higher than those from 57507905 and 56463 (Group 1). There are high similarities of atoms pairs from 57507911 (Group 1), on the one hand, and 44397540 and 11616723 (Group 2), on the other hand (higher than those with other molecules from Group 1). Similarities above 0.60, which are significant, are noticed for other pairs of molecules belonging to different groups.

A similar result has been obtained using the Tanimoto coefficient for molecules' similarities. Table 7 displays the Tanimoto coefficients and their ranked similarities (from 1 to 7) for pairs of molecules: one from the first group and another from the second group. Blank means a similarity rank over 8. There is only one molecule not listed in Table 7(ID 53394099, Group 2), whose similarity with all the molecules in Group 1 is very low.

Based on the structural characteristics, the molecules' clustering has been performed. The results for Group 1 are presented in the following. The binning clustering, with cutoffs (lower similarity bounds) of 0.3, 0.6, and 0.9 resulted in 1, 2, or 7 clusters, respectively (for both groups). The Jarvis–Patrick algorithm provided 2 clusters when considering four neighbors, and a single cluster when considering 5 and 6 neighbors. The optimal number of clusters (3) in the k-means algorithm has been selected by the elbow method.

Running the k-means algorithm resulted in three clusters with the sizes 1, 3, and 3 and between the sum of squares/total sum of squares = 77.0%. The first clusters contain the first three molecules; the second one, the next three molecules; and the last cluster contains molecule CID 56463.

For the second group, the algorithms binning and Jarvis–Patrick algorithms provided a single cluster.

The heat map and the dendrogram resulted from the hierarchical clustering are presented in Figure 6 for Group 1 and in Figure 7 for Group 2.

In the heat maps, the darker the color is, the lower the distance between the compounds is, so the higher the similarity is. Hence, the pair of compounds whose corresponding squares are dark green are the most similar. A very low similarity is noticed for the pairs of compounds situated in the bottom right-hand side of both heat maps.
The dendrograms (right-hand side of Figures 6 and 7) confirm the similarity strength. The higher the branch between two compounds is, the lower the similarity is.

In Figure 6, the highest similarity is found between the groups of molecules (117587582 and 9889172) and (79507911 and 56841530). They both have a Tanimoto coefficient of 0.900 and an overlap coefficient of 0.9583 (listed in Table 4 for the first pair, not listed for the second one for the lack of space). From the viewpoint of similarity intensity, the second place is occupied by the pairs of molecules (10226340, 117587582) and (10226340, 9889172), with a Tanimoto coefficient of 0.8519 and the same overlap coefficient.

In Figure 7, the most similar molecules are in the pairs (44397500, 11616723), (44397540, 11616723), and (11616723, 11493740), with a Tanimoto coefficient of 0.9152 and the overlap coefficient 0.9011. They are not listed in the tables to avoid repeating similar information.

Analyzing Figures 6 and 7, remark that the molecules in Group 2 have the distance between them, measured on the dendrograms, lower than those from Group 1, confirming the previous findings related to the molecules’ similarity. The same procedures have been utilized, putting together all the 14 molecules. The heat map and the dendrogram are shown in Figure 8.

The highest similarities are found for the triple (44397500, 44397540, 11616723) from Group 2 and the pairs (957507905, 56481530) and (117587582, 9889172) from the first group.
Figure 5: The structure of the molecule 9889172.

Table 1: The code identification (CID), molecular formula (MF), molar weight filter (MW), types, and the numbers of atoms, rings, and aromatic rings in the molecules.

| CID           | Formula        | MW     | C  | F  | H  | I  | N  | O  | Cl | Rings | Aromatic |
|---------------|----------------|--------|----|----|----|----|----|----|----|-------|----------|
| 9889172       | C20H19FINO     | 435.2738 | 20 | 1  | 19 | 1  | 1  | 0  | 3  | 3     |          |
| 117587582     | C20H19CINO     | 451.7284 | 20 | 0  | 19 | 1  | 1  | 1  | 1  | 3     | 3        |
| 10226340      | C22H23IN2O     | 458.3353 | 22 | 0  | 23 | 1  | 2  | 1  | 0  | 4     | 3        |
| 56841530      | C21H23NO2      | 321.4128 | 21 | 0  | 23 | 0  | 1  | 2  | 0  | 3     | 3        |
| 57507911      | C21H23NO2      | 321.4128 | 21 | 0  | 23 | 0  | 1  | 2  | 0  | 3     | 3        |
| 57507905      | C20H21NO2      | 307.3862 | 20 | 0  | 21 | 0  | 1  | 2  | 0  | 3     | 3        |
| 56463         | C23H26N2O3     | 378.4641 | 23 | 0  | 26 | 0  | 2  | 3  | 0  | 4     | 3        |

Group 1

| 44397641      | C22H25NO2      | 335.4394 | 22 | 0  | 25 | 0  | 1  | 2  | 0  | 3     | 3        |
| 44397500      | C21H22CINO     | 339.8585 | 21 | 0  | 22 | 0  | 1  | 1  | 1  | 3     | 3        |
| 44397540      | C22H25NO2      | 335.4394 | 22 | 0  | 25 | 0  | 1  | 2  | 0  | 3     | 3        |
| 53494930      | C24H28N2O2     | 376.4913 | 24 | 0  | 28 | 0  | 2  | 2  | 0  | 4     | 3        |
| 11616723      | C22H25NO2      | 319.4400 | 22 | 0  | 25 | 0  | 1  | 1  | 0  | 3     | 3        |
| 11493740      | C22H25NO2      | 335.4394 | 22 | 0  | 25 | 0  | 1  | 2  | 0  | 3     | 3        |
| 53394099      | C25H29NO2      | 375.5033 | 25 | 0  | 29 | 0  | 1  | 2  | 0  | 4     | 3        |

Group 2

Table 2: Molecules’ descriptors.

| CID           | HBA1 | HBA2 | HBD | logP  | MR    | TPSA |
|---------------|------|------|-----|-------|-------|------|
| 9889172       | 20   | 2    | 0   | 5.6167| 105.0705| 22.00|
| 117587582     | 20   | 2    | 0   | 5.8860| 109.8155| 22.00|
| 10226340      | 25   | 3    | 0   | 4.8991| 119.3305| 25.24|
| 56841530      | 25   | 3    | 0   | 5.0711| 98.7945 | 31.23|
| 57507911      | 25   | 3    | 0   | 5.0711| 98.7945 | 31.23|
| 57507905      | 23   | 3    | 0   | 4.6810| 93.9875 | 31.23|
| 56463         | 30   | 4    | 0   | 3.5954| 115.3485| 43.70|

Group 1

| 44397641      | 27   | 3    | 0   | 5.2655| 103.6015| 31.23|
| 44397500      | 23   | 2    | 0   | 5.9103| 102.1195| 22.00|
| 44397540      | 27   | 3    | 0   | 5.2655| 103.6015| 31.23|
| 53494930      | 31   | 4    | 0   | 4.4975| 117.9125| 34.47|
| 11616723      | 26   | 2    | 0   | 5.5653| 102.0755| 22.00|
| 11493740      | 27   | 3    | 0   | 5.2655| 103.6015| 31.23|
| 53394099      | 31   | 3    | 0   | 6.0457| 115.9085| 31.23|
Table 3: Tanimoto coefficients for the atoms similarity.

| CID          | 9889172 | 117587582 | 10226340 | 57507911 | 56841530 | 57507905 | 56463 |
|--------------|---------|-----------|----------|----------|----------|----------|-------|
| 9889172      | 1.0000  | 0.8462    | 0.6788   | 0.6140   | 0.5376   | 0.4901   | 0.4469|
| 117587582    | 1.0000  | 0.6788    | 0.5376   | 0.6140   | 0.4901   | 0.4469   |       |
| 10226340     | 1.0000  | 0.3139    | 0.5774   | 0.4859   | 0.5054   |          |       |
| 57507911     | 1.0000  | 0.7250    | 0.8467   | 0.6802   | 0.5545   |          |       |
| 56841530     | 1.0000  | 0.4953    |          |          |          |          |       |
| 57507905     | 1.0000  | 0.4469    |          |          |          |          |       |
| 56463        | 1.0000  | 0.4901    |          |          |          |          |       |

Table 4: Similarity measures for the molecule CID 9889172 and the molecules in Group 1.

| CID          | Query_Size | Target_Size | Tanimoto_Coefficient | Overlap_Coefficient |
|--------------|------------|-------------|----------------------|---------------------|
| 9889172      | 24         | 24          | 1.0000               | 1.0000              |
| 117587582    | 24         | 24          | 0.9200               | 0.9583              |
| 10226340     | 24         | 26          | 0.8519               | 0.9583              |
| 57507911     | 24         | 24          | 0.8469               | 0.9167              |
| 56841530     | 24         | 24          | 0.8469               | 0.9167              |
| 57507905     | 24         | 23          | 0.8077               | 0.9130              |
| 56463        | 24         | 28          | 0.5758               | 0.7917              |

Table 5: Similarity measures for the molecule CID 44397641 and the molecules in Group 2.

| CID          | Query_Size | Target_Size | Tanimoto_Coefficient | Overlap_Coefficient |
|--------------|------------|-------------|----------------------|---------------------|
| 44397641     | 25         | 25          | 1.0000               | 1.0000              |
| 44397500     | 25         | 24          | 0.8846               | 0.9583              |
| 44397540     | 25         | 25          | 0.8519               | 0.9200              |
| 53494930     | 25         | 28          | 0.7667               | 0.9200              |
| 11616723     | 25         | 24          | 0.8846               | 0.9583              |
| 11493740     | 25         | 25          | 0.8519               | 0.9200              |
| 53394099     | 25         | 28          | 0.7667               | 0.9200              |

Table 6: Pair-atoms similarities for molecules from different groups.

| CID          | 44397641 | 44397500 | 44397540 | 53494930 | 11616723 | 11493740 | 53394099 |
|--------------|----------|----------|----------|----------|----------|----------|----------|
| 9889172      | 0.4883   | 0.5248   | 0.4922   | 0.4566   | 0.5248   | 0.4884   | 0.4598   |
| 117587582    | 0.4883   | 0.5462   | 0.4922   | 0.4566   | 0.5248   | 0.4884   | 0.4598   |
| 10226340     | 0.4706   | 0.4913   | 0.4810   | 0.6087   | 0.5062   | 0.4741   | 0.4894   |
| 57507911     | 0.6364   | 0.5817   | 0.6842   | 0.5646   | 0.6677   | 0.6600   | 0.5000   |
| 56841530     | 0.6649   | 0.5249   | 0.6134   | 0.5139   | 0.5726   | 0.7194   | 0.4696   |
| 57507905     | 0.6029   | 0.4901   | 0.5937   | 0.4847   | 0.5333   | 0.6361   | 0.4244   |
| 56463        | 0.4803   | 0.3915   | 0.4456   | 0.4595   | 0.42173  | 0.4934   | 0.4766   |

Table 7: Tanimoto coefficients for the atoms similarity in different groups of molecules.

| CID          | 44397641 | 44397500 | 44397540 | 53494930 | 11616723 | 11493740 | 53394099 |
|--------------|----------|----------|----------|----------|----------|----------|----------|
| 9889172      | 0.5262   | (6)      |          | 0.5262   | (7)      |          |          |
| 117587582    |          |          |          | 0.5262   | (7)      |          |          |
| 10226340     |          |          |          | 0.6096   | (4)      | 0.5075   | (7)      |
| 57507911     | 0.6374   | (7)      | 0.6851   | (3)      |          |          | 0.6609   (6) |
| 56841530     | 0.6657   | (5)      | 0.6145   | (6)      | 0.7202   | (4)      |          |
| 57507905     | 0.6040   | (5)      | 0.5948   | (6)      | 0.6371   | (4)      |          |
| 56463        | 0.4815   | (7)      |          |          | 0.4945   | (6)      |          |
The clusters determined by the k-means algorithm on the 14 molecules are shown in Figure 9.

The first cluster contains all the molecules in Group 2, a result consistent with the previous findings. The first group is divided into two clusters: the first one containing the first three molecules and the second one, the rest. The result is consistent with the dendrogram in Figure 6. Cluster 2 (Figure 9) contains the same elements as the top subtree in Figure 7.
4. Conclusions

Cheminformatics is a study field that offers many facilities in collecting, organizing, and studying chemical information [45]. Its contribution is essential to drug design [46, 47]. In this view, our article focused on some applications of Cheminformatics to analyze the benzoylindoles and phenylacetylindoles groups of drugs. The structures and molecules functional groups have been emphasized using the facilities provided by rdk and ChemmineR packages. Structural similarities of different molecules have been determined using the Tanimoto index and clustering techniques. A higher similarity in the second group compared to the first one is observed. Further practical studies should confirm the hypothesis of the similar actions and effects of these drugs and the possible cure using the same inhibitors.

Data Availability

Data can be freely downloaded from https://pubchem.ncbi.nlm.nih.gov/ by introducing the compound name in the search window.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

References

[1] S. O’Mahony Carey, “Psychoactive Substances. A Guide to Ethnobotanical Plants and Herbs, Synthetic Chemicals, Compounds and Products”, http://lab.bnn.go.id/nps_alert_system/publikasi%20web/Psychoactive%20plant/Psychoactive_plant.pdf.

[2] J. van Amsterdam, D. Nutt, and W. van den Brink, “Generic legislation of new psychoactive drugs,” *Journal of Psychopharmacology*, vol. 27, no. 3, pp. 317–324, 2013.

[3] C. Vlădescu, S. G. Scîntee, V. Olsavszky, C. Hernández-Quevedo, and A. Sagan, “Romania: health system review,” *Health Systems in Transition*, vol. 18, no. 4, pp. 1–170, 2016.

[4] L. Elliott, C. K. Haddock, S. Campos, and E. Benoit, “Poly-substance use patterns and novel synthetics: a cluster analysis from three U.S. cities,” *PLoS One*, vol. 14, no. 12, Article ID e0225273, 2019.
[5] A. Hardon, “Chemical breath,” Critical Studies in Risk and Uncertainty, vol. 6, pp. 81–111, 2021.
[6] A. Shafi, A. J. Berry, H. Sumnall, D. M. Wood, and D. K. Tracy, “New psychoactive substances: a review and updates,” Therapeutic advances in psychopharmacology, vol. 10, Article ID 2045125320967197, 2020.
[7] EMCCDA 2016, European Monitoring Center for Drugs and Drug Addiction, “Perspectives on Drugs Health Responses to New Psychoactive Substances,” https://www.emccda.europa.eu/system/files/publications/2933/NPS%20health%20respond es.Pod2016.pdf.
[8] EMCCDA 2017, European monitoring center for drugs and drug addiction, “perspectives on drugs synthetic cannabinoids in europe” https://www.emccda.europa.eu/system/files/p ublications/2753/POD_Synthetic%20cannabinoids_0.pdf.
[9] L. Ernst, N. Langer, A. Bockelmann, E. Salkhordeh, and T. Beuerle, “Identification and quantification of synthetic cannabinoids in ‘spice-like’ herbal mixtures: update of the German situation in summer 2018,” Forensic Science International, vol. 294, pp. 96–102, 2019.
[10] A. D. Lesiak and J. R. Shepard, “Recent advances in forensic drug analysis by DART-MS,” Bioanalysis, vol. 6, no. 6, pp. 819–842, 2014.
[11] S. Mignani, J. Rodrigues, H. Tomas et al., “Present drug-likeness filters in medicinal chemistry during the hit and lead optimization process: how far can they be simplified?” Drug Discovery Today, vol. 23, no. 3, pp. 605–615, 2018.
[12] P. J. Rogers, “Food and drug addictions: similarities and differences,” Pharmacology Biochemistry and Behavior, vol. 153, pp. 182–190, 2017.
[13] F. Zapata, J. M. Matey, G. Montalvo, and C. García-Ruiz, “Chemical classification of new psychoactive substances (NPS),” Microchemical Journal, vol. 163, Article ID 105877, 2021.
[14] V. L. Alves, J. L. Gonçalves, J. Aguiar, H. M. Teixeira, and J. S. Câmara, “The synthetic cannabinoids phenomenon: from structure to toxicological properties. A review,” Critical Reviews in Toxicology, vol. 50, no. 5, pp. 359–382, 2020.
[15] K. Soltaninejad, “Clinical and forensic toxicological aspects of synthetic cannabinoids: a review and update,” Asia Pacific Journal of Medical Toxicology, vol. 9, no. 3, pp. 108–118, 2020.
[16] A. J. Potts, C. Cano, S. H. L. Thomas, and S. L. Hill, “Synthetic cannabinoid receptor agonists: classification and nomenclature,” Clinical Toxicology, vol. 58, no. 2, pp. 82–98, 2020.
[17] J. W. Godden, F. L. Stahura, and J. Bajorath, “Anatomy of fingerprint search calculations on structurally diverse sets of active compounds,” Journal of Chemical Information and Modeling, vol. 45, no. 6, pp. 1812–1819, 2005.
[18] F. L. Dragomir, “Modeling resource in E-commerce,” in Proceedings of the 10th International Conference on Knowledge Management: Projects, Systems and Technologies, Bucharest, Romania, November 2017.
[19] F. L. Dragomir, “Models of digital markets,” in Proceedings of the 10th International Conference on Knowledge Management: Projects, Systems and Technologies, Bucharest, Romania, November 2017.
[20] F. L. Dragomir, “Approaches to modeling consultancy systems by recommendation,” in Proceedings of the 10th International Conference on Knowledge Management: Projects, Systems and Technologies, Bucharest, Romania, November 2017.
[21] F. L. Dragomir and G. Alexandrescu, “Operational Research–Theoretical Bases,” Sitech, Craiova, Romania, 2017, in Romanian.
[22] A. Voicu, N. Duteanu, M. Voicu, D. Vlad, and V. Dumitrescu, “The rcdk and cluster R packages applied to drug candidate selection,” Journal of Cheminformatics, vol. 12, p. 3, 2020.
[23] R. Iswandana, P. Aisyah, and R. R. Syahidi, “Prediction analysis of pharmacokinetic parameters of several oral systemic drugs using insilico method,” International Journal of Applied Pharmacaceutics, vol. 12, no. 1, pp. 260–263, 2020.
[24] S. P. Leelananda and S. Lindert, “Computational methods in drug discovery,” Beilstein Journal of Organic Chemistry, vol. 12, pp. 2694–2718, 2016.
[25] G. Sliwoski, S. Kothiwale, J. Meiler, and E. W. Lowe Jr., “Computational methods in drug discovery,” Pharmacological Reviews, vol. 66, no. 1, pp. 334–395, 2014.
[26] P. Willett, “Similarity searching using 2D structural fingerprints,” Methods in Molecular Biology (Clifton, N.J.), vol. 672, pp. 133–158, 2011.
[27] R. Guha, K. Gilbert, G. Fox, M. Pierce, D. Wild, and H. Yuan, “Advances in cheminformatics methodologies and infrastructure to support the data mining of large, heterogeneous chemical datasets,” Current Computer-Aided Drug Design, vol. 6, no. 1, pp. 50–67, 2010.
[28] Y. Cao, A. Charisi, L.-C. Cheng, T. Jiang, and T. Girke, “ChemmineR: a compound mining framework for R,” Bioinformatics, vol. 24, no. 15, pp. 1733-1734, 2008.
[29] Y. Wang, T. W. H. Backman, K. Horan, and T. Girke, “RmscR: mismatch tolerant maximum common substructure searching in R,” Bioinformatics, vol. 29, no. 21, pp. 2792–2794, 2013.
[30] R. Guha, “Chemical informatics functionality in R,” Journal of Statistical Software, vol. 18, no. 5, pp. 1–16, 2007.
[31] R. Guha and M. R. Cherto, “Integrating the CDK with R,” Chemical informatics functionality in R, vol. 1–17, 2017, http://cran.nexr.com/web/packages/rcdk/vignettes/rcdk.pdf.
[32] S. Mente and M. Kuhn, “The use of the R language for medicinal chemistry applications,” Current Topics in Medicinal Chemistry, vol. 12, no. 18, p. 1957, 2012.
[33] PubChem. Available at: https://pubchem.ncbi.nlm.nih.gov.
[34] H. Kubinyi, “Hydrogen bonding: the last mystery in drug design?,” in Pharmacokinetic Optimization in Drug Research: Biological, Physicochemical, and Computational Strategies, B. Testa, H. van de Waterbeemd, G. Folkers, and R. Guy, Eds., pp. 513–521, Verlag Helvetica Chimica Acta, Zürich, Switzerland, 2001.
[35] G. Caron, M. Vallaro, and G. Ermondi, “Log P as a tool in intramolecular hydrogen bond considerations,” Drug Discovery Today: Technologies, vol. 27, pp. 65–70, 2018.
[36] G. L. Patrick, An Introduction to Medicinal Chemistry, Oxford University Press, Oxford, UK, 1995.
[37] S. A. Cuesta, J. R. Mora, and E. A. Marquez, “In silico classification and recognition forces and property spaces,” Current Computer-Aided Drug Design, vol. 12, no. 20, pp. 133–158, 2016.
[38] P. Ertl, B. Rohde, and P. Selzer, “Fast calculation of molecular fingerprints on structurally diverse sets of active compounds,” Journal of Chemical Information and Modeling, vol. 4, no. 6, pp. 1812–1819, 2005.
[39] G. Caron, M. Vallaro, and G. Ermondi, “Log P as a tool in intramolecular hydrogen bond considerations,” Drug Discovery Today: Technologies, vol. 27, pp. 65–70, 2018.
[40] J. V. Turner and S. Agatonovic-Kustrin, “In silico prediction of oral bioavailability,” Comprehensive Medicinal Chemistry II, vol. 5, pp. 699–724, 2007.
[41] D. Bajusz, A. Rácz, and K. Héberger, "Why is Tanimoto index an appropriate choice for fingerprint-based similarity calculations?" *Journal of Cheminformatics*, vol. 7, no. 1, 2015.

[42] A. Tversky, "Features of similarity," *Psychological Review*, vol. 84, no. 4, pp. 327–352, 1977.

[43] F. Murtagh and P. Contreras, "Algorithms for hierarchical clustering: an overview," *Wiley Interdisciplinary Reviews. Data Mining Knowledge Discovery*, vol. 2, no. 1, pp. 86–97, 2012.

[44] R. A. Jarvis and E. A. Patrick, "Clustering using a similarity measure based on shared near neighbors," *IEEE Transactions on Computers*, vol. C-22, no. 11, pp. 1025–1034, 1973.

[45] M. Karthikeyan and R. Vyas, "Machine learning methods in chemoinformatics for drug discovery," in *Practical Chemoinformatics*, M. Karthikeyan and R. Vyas, Eds., Springer, New Delhi, India, pp. 133–194, 2014.

[46] J. A. DiMasi, R. W. Hansen, and H. G. Grabowski, "The price of innovation: new estimates of drug development costs," *Journal of Health Economics*, vol. 22, no. 2, pp. 151–185, 2003.

[47] M. Hassan Baig, K. Ahmad, S. Roy et al., "Computer aided drug design: success and limitations," *Current Pharmaceutical Design*, vol. 22, no. 5, pp. 572–581, 2016.