Research Article GU J Sci 38(3): 1480-1501 (2025) DOI: 10.35378/gujs.1581396

Gazi University

Journal of Science

http://dergipark.org.tr/gujs

Evaluations of Comet Assay Data Through Statistical Analysis, Machine Learning, and Multi-Criteria Decision-Making Methods for Genotoxic Potential of Food Sweeteners

Fatma UNAL[1,*], Muhlis OZDEMIR[2], Deniz YUZBASIOGLU[1], Sevcan MAMUR[3], Ece AVULOGLU YILMAZ[4], Fatma OKUS[5], Ece AKBAS[1], Resat KASAP[2], Busra GUNES[5]

[1]*Department of Biology, Science Faculty, Gazi University, 06500, Ankara, Türkiye*
[2]*Department of Statistics, Science Faculty, Gazi University, 06500, Ankara, Türkiye*
[3]*Department of Environmental Sciences, Graduate School of Natural and Applied Sciences, Gazi University, 06560, Ankara, Türkiye*
[4]*Department of Computer Technologies, Vocational School of Technical Sciences, Amasya University, 05000, Amasya, Türkiye*
[5]*Department of Biology, Graduate School of Natural and Applied Sciences, Gazi University, 06560, Ankara, Türkiye*

Highlights

- D-sorbitol-DS, Xylitol-XYL, and mixture-DSX are widely used as food sweeteners.
- Increasing in comet tail length, intensity, and moment is an indication of DNA damage.
- ML and MCDM methods enabled a more straightforward interpretation.
- They provided a more comprehensive comparison of the genotoxic effects of sweeteners.
- Most genotoxic sweeteners are XYL, DSX, and D-Sorbitol, respectively.

Article Info

Received:12 Dec 2024
Accepted: 20 May 2025

Keywords

Comet assay
DNA damage
Statistical analysis
Machine learning
Multi-criteria decision-making

Abstract

Although the genotoxic effects of food sweeteners have been studied previously, there is still a lack of application using an integrated approach that combines statistical analysis, Machine Learning (ML), and Multi-Criteria Decision Methods (MCDM) in depth to reveal the DNA-damaging potential of food sweeteners (D-Sorbitol (DS) and Xylitol (XYL)), both alone and in combination (DSX). A dataset on comet assay observations for DNA damage (tail length, tail intensity, and tail moment) was collected from previous studies. Kruskal-Wallis and One-Way ANOVA tests were used to identify significant differences in DNA damage. K-Means and Hierarchical Clustering lead to grouping additives of genotoxic effects, while MOORA and TOPSIS ranked toxicity levels. The findings of MCDM showed that XYL_1000 caused the highest DNA damage (0.726683 and 0.382296). The combination of DS and XYL (DSX_M8) exhibited higher toxicity (0.715258 and 0.37281) compared to their treatments, whereas DSX M1 revealed the least damaging effect (0.235688 and 0.0324946). This is the first study using this approach. These findings highlight the impact of combining ML and MCDM methods for a more intensive genotoxicity evaluation, providing precious insights into food safety regulations.

1. INTRODUCTION

Food additives are chemicals added to food to keep and enhance its nutrition, flavor, appearance, color, and flexibility in the food industry [1]. One of the food additives is artificial sweeteners, which serve as a sugar substitute, being sweeter than table sugar and with fewer calories, improving taste and pleasure in various food products [2,3]. Sweeteners are classified as sugar alcohols (polyols), natural sweeteners, and artificial sweeteners [4,5]. Polyols are naturally occurring carbohydrates found in small quantities in fruits and vegetables and are obtained by replacing an aldehyde or keto group with a hydroxyl group [6,7]. Sorbitol (E420), also known as D-Sorbitol (DS), is a six-carbon polyol that is a white and crystalline compound that is about 40-70% as sweet as sucrose and is used in jams, chocolates, candies, biscuits, cereal bars, and soft drinks [8-10]. Xylitol (E967) (XYL) is the most frequently used five-carbon crystalline polyol that is almost as sweet as sucrose and is used in peppermint, chewing gum, lozenges, and other confectionery and dental

*Corresponding author, e-mail: funal@gazi.edu.tr

and pharmaceutical products [2, 4, 11-13]. While artificial sweeteners are approved for consumption by the FDA and used in many areas of daily life, there are some concerns about their side effects, especially genotoxic effects after long-term use of such chemicals [3, 14-16]. The genotoxic effects of chemicals can be determined using various short-term tests. One of them, the comet or single-cell gel electrophoresis (SCGE) assay, is a versatile, sensitive, efficient, and easy to evaluate multiple DNA lesions simultaneously by the statistical comparison of differential electrophoretic migration in various cell types [17-19]. DNA damage is a valuable biomarker due to its role in cancer and age-related diseases such as cardiovascular, diabetes, Alzheimer's, and Parkinson's disease [20-23].

This article describes the development of a novel predictive model for evaluating the genotoxicity of food sweeteners using MCDM methods, MOORA and TOPSIS, and ML approaches. Factors such as variations in assessments applied across studies, differences in treatment ways and cell types, and chemicals produced make it difficult for scientists to analyze and effectively compare the toxicities of chemicals. On the other hand, computational approaches for predicting genotoxicity based on chemical structures and properties are recognized as an alternative to experimental analysis [24]. Therefore, machine learning (ML) techniques are applied as a powerful tool for predicting toxicity based on available data.

In the literature, no study "evaluating" the genotoxicity of food sweeteners using machine learning techniques as we applied in our model. This diversity highlights the novelty and originality of this study. However, there are some investigations for "prediction" of cellular responses and genotoxic effects, specifically for nanoparticles, using various ML techniques because conventional experimental procedures for evaluating nanoparticle toxicity are expensive and time-consuming. The training dataset used to establish these models includes the physicochemical characteristics of nanoparticles, treatment conditions, and cellular responses of different cell lines. In one of these publications, Sizochenko et al. [25] evaluated the genotoxicity of 49 silicon and metal oxide nanoparticles in Ames and Comet tests using supervised quantitative structure-activity relationship (QSAR) based on classification modeling techniques. They verified that *the in vitro* Comet assay is helpful for fast screening tests, as it reproduces *in vivo* conditions better than the Ames test. They also stated that the Ames test is unsuitable for fast nanoparticle genotoxicity/mutagenicity screening. Xu et al. [26] predicted the cytotoxicity of micro- and nanoplastics by combining machine learning with meta-analysis. The results obtained from 34 published articles on the relationship between micro- and nanoplastics and toxicity using various cell models were evaluated using the decision tree and decision tree ensemble method for modeling the cytotoxicity. They reported the decision tree ensemble (DTE1) as the best prediction model. Another study applied machine learning models, including linear regression, decision trees, neural networks, and instance-based algorithms to predict the cytotoxicity of nanomaterials. It has provided insight into predicting the potential of nanomaterial toxicity by developing multiple models using various machine-learning algorithms. Applying machine learning techniques in this way holds promise for nanoscience and nanotechnology [27]. Fan et al. [24] carried out *in silico* prediction of chemical genotoxicity using the classification models of six machine learning methods involving the support vector machine (SVM), naïve Bayes (NB), k-nearest neighbor (kNN), C4.5 decision tree (DT), random forest (RF) and artificial neural network (ANN). Authors suggested that these models might be helpful for the initial screening of potentially genotoxic compounds. ML techniques also determined the classification of segmented comet assay images [28,29]. Atila et al. [28] classified DNA damages of comet assay images using a convolutional neural network. Hong et al. [29] analyzed the classification of comet segmentation using the deep learning method. They demonstrated that the newly proposed DL-based comet assay analysis program, DeepComet, has performed excellently. Based on such investigations, we have applied an in-depth integrated approach that combines statistical analysis, Machine Learning (ML), and Multi-Criteria Decision Methods (MCDM) to evaluate the DNA-damaging potential of food sweeteners.

Genotoxicity studies cause concerns due to costly and ethically questionable *in vivo* tests on animals and *in vitro* experiments to assess their bioactivity. Experimental analysis is also generally expensive and time-consuming. Such factors increase the importance of the detection and ranking of the toxicity potential of chemical compounds [30]. Recently, technologies applying computational methods have become increasingly crucial in various scientific areas. Applying machine learning (ML) to evaluate DNA damage by ML [31,32] may represent an empirical approach in genotoxicity studies using the comet assay. ML is

a set of techniques recognized as a subset of artificial intelligence (AI) that enables fast and correct conclusions for new and diverse situations by learning from the data calculated in a specific area. ML applications have been used in many scientific fields, including cancer, genetic toxicology, and image processing [24,31,32]. Machine learning techniques have become increasingly critical in predicting the genotoxicity of compounds, as they can handle large datasets with complex, non-linear relationships. Algorithms such as Support Vector Machines (SVM), Random Forests (RF), and k-nearest Neighbors (k-NN) have demonstrated strong performance in predicting DNA damage induced by chemicals. A study employing ML models found that SVM and RF achieved 80-90% accuracy rates in classifying genotoxic and non-genotoxic compounds, offering a more scalable and precise alternative to traditional methods [33,34]. Additionally, Multi-Criteria Decision-Making (MCDM) methods such as MOORA and TOPSIS are increasingly being integrated into genotoxicity studies to rank and compare the effects of different substances based on multiple criteria. These approaches allow for considering several factors simultaneously, providing a more holistic evaluation of a compound's genotoxicity. While the application of MCDM in this field remains limited, studies suggest that it could significantly improve decision-making by providing more explicit guidance on the relative risks associated with various substances [35,36].

All the DNA damages determined by comet assay parameters, tail length (μm), tail intensity (DNA%), and tail moments have a numerical value to perform quantitative methods like statistical analysis to compare groups whether they are statistically significant, ML for clustering and multi-criteria decision making (MCDM) for ranking. Therefore, this study aimed to apply three different approaches, statistical analysis, ML, and MCDM, to reveal and evaluate the genotoxic potential of two food sweeteners, one by one, and their combination when used as a mix. In this study, genotoxic effects of DS, XYL, and DS+XYL (Mix, DSX) determined by the comet assay carried out in our laboratory were evaluated using Kruskall Wallis, One-Way ANOVA, K-Means and Hierarchical Clustering and MCDM Methods, including MOORA (The Multi-Objective Optimization Ratio Analysis) and TOPSIS (Technique for Order Preference by Similarity to Ideal Solution). While standard comet assay evaluates damages in terms of DNA tail length (μm), tail intensity (DNA%), and tail moment, individually, for each test chemical or their mixture, the methods used in the present study evaluated DNA damages using all of these three parameters in a single application and in a shorter of time. Moreover, six techniques were applied to compare the results obtained from the experimental and control groups in the present study over the standard comet assay, where only one statistical analysis (z-test) was used. Therefore, the innovative aspect of this study is to be the first investigation in which the genotoxic effects of sweeteners administered both alone and in combination to human lymphocytes were evaluated using various statistical analyses, ML approaches, and MCDM methods that grouped and ranked all the treatments instead of traditional statistical methods.

2. MATERIAL AND METHODS

2.1. Application of Alkaline Comet Assay

The comet assay introduced by Singh et al. [37] was applied with some modifications [38,39]. Avuloglu-Yilmaz et al. [4] and Gunes [40] give details of the experimental studies. Both studies were carried out with the permission of the Gazi University Clinical Research Ethics Committee (11.06.2018/452 and 04.04.2022-261 / 05.12.2022-886, respectively). Isolated lymphocytes were treated with various concentrations of food sweeteners for one hour at 37°C. A positive (100 μM hydrogen peroxide-H_2O_2) and a negative control (distilled water) were also run (Table 1). For each concentration, 200 randomly selected cells (comets) were analyzed in terms of tail length (μm), tail intensity, and tail moment using a specialized image analysis system (Comet Assay IV, Perceptive Instruments Ltd., UK) [41,42].

2.2. Data Obtaining

The raw data used in this study were obtained during experiments carried out by previous publications in our laboratory. In these experiments, two different sweeteners [(D-Sorbitol (DS) and Xylitol (XYL)] and their mixture (DSX) [4, 40] were used. The data were obtained from 200 cells from 2 different donors (100 per person). While five and four different concentrations were used for DS and XYL, respectively, eight

various combinations were applied for DSX. In addition, each treatment had a positive and a negative control group. Table 1 shows the concentrations and abbreviations used in this study.

Table 1. *Sweeteners, abbreviations, concentrations, and number of observations used for analysis*

Sweeteners	**Concentration (µg/mL)**		**Abbreviations**	**# of observations for comet tail length, tail intensity, and tail moment (200 for each)**
D Sorbitol (DS)	Negative Control		DS_C	200
	Positive control (H_2O_2-100 mM)		DS_C+	200
	62.5		DS_62.5	200
	125		DS_125	200
	250		DS_250	200
	500		DS_500	200
	1000		DS_1000	200
Xylitol (XYL)	Negative Control		XYL_C	200
	Positive control (H_2O_2-100 mM)		XYL_C+	200
	125		XYL_125	200
	250		XYL_250	200
	500		XYL_500	200
	1000		XYL_1000	200
Mixture (M)	**D-Sorbitol + (volume)**	**Xylitol (volume)**		
	Negative Control		DSX_C	200
	Positive control (H_2O_2-100 mM)		DSX_C+	200
	500 (1)	500 (1)	DSX_M1	200
	500 (1)	1000 (1)	DSX_M2	200
	1000 (1)	500 (1)	DSX_M3	200
	1000 (1)	1000 (1)	DSX_M4	200
	500 (3)	500 (1)	DSX_M5	200
	500 (3)	1000 (1)	DSX_M6	200
	1000 (3)	500 (1)	DSX_M7	200
	1000 (3)	1000 (1)	DSX_M8	200
# of tail length				4600
# of tail intensity				4600
# of tail moment				4600
In total				**13800**

2.3. Methodology

This section consists of three subsections. The first subsection involves the application of statistical methods, the second subsection focuses on applying ML methods, and the final subsection applies the MCDM methods. Kruskal Wallis and One-Way ANOVA tests were used in this study to identify differences among food additives and reveal the DNA damage in cells induced by food sweeteners. K-means and hierarchical clustering, two ML methods, were employed to cluster the food sweeteners according to cell DNA damage and their similarities and dissimilarities. Finally, two different MCDM methods, MOORA and TOPSIS, were applied to rank and compare food sweeteners according to their hazardous effects on cell DNA. Figure 1 summarizes the processes carried out in this study as follows.

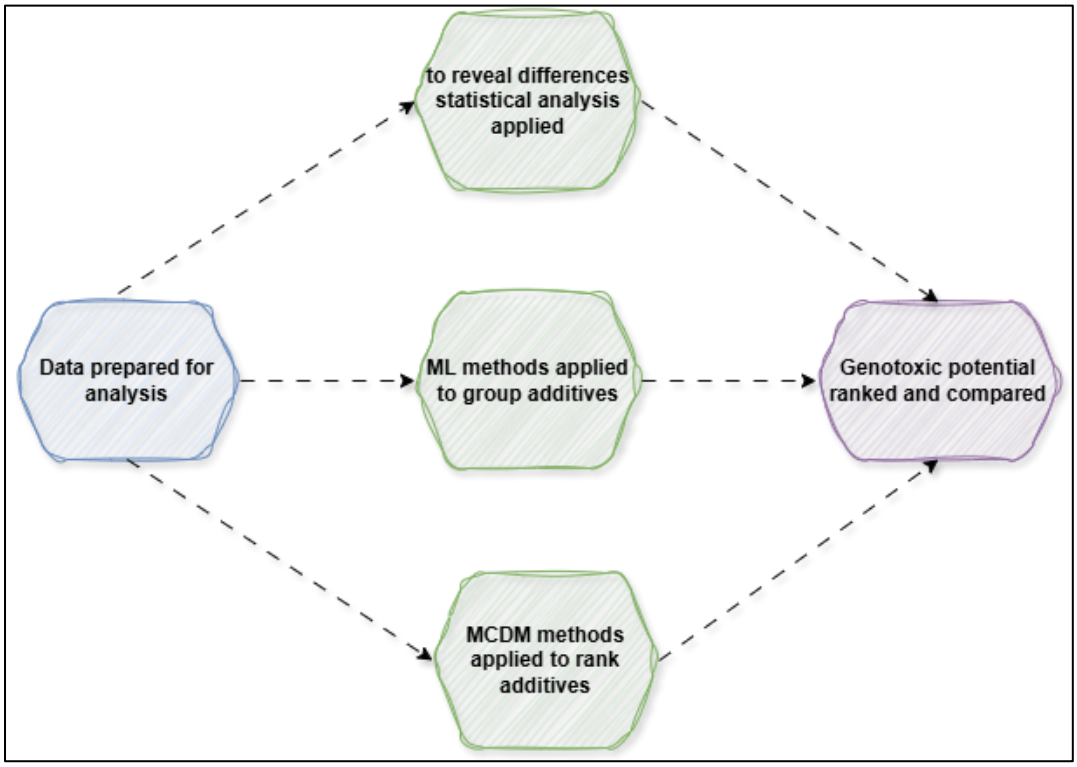


Figure 1. *Workflow for evaluating the genotoxic potential of food additives using statistical analysis, ML, and MCDM methods*

2.3.1. Kruskal Wallis and One-Way ANOVA

Since normality and equal variance assumptions were not met, the Kruskal-Wallis test was used to compare comet tail length, tail intensity, and tail moment across groups. One-way ANOVA was conducted to analyze these differences further, revealing significant variation in mean scores for tail length, tail intensity, and tail moment across food additive groups. IBM SPSS Statistics (Version 25) was used to perform analysis for Kruskal Wallis and One-Way ANOVA.

2.3.2. K-Means and hierarchical clustering

K-means and hierarchical clustering are widely used, and popular ML methods are briefly explained below. The preferred approaches used for each method are detailed below. The programming language and its packages used to implement these methods are discussed thereafter. Both methods are unsupervised clustering methods in ML. This study applied K-means and hierarchical clustering to cluster sweeteners according to their effects on cell DNA. The difference between the two methods can be explained as follows. Although K-means clustering is one of the most preferred techniques in ML since it is a straightforward and fast algorithm, it requires pre-specifying the K as the number of clusters. The hierarchical clustering approach does not need to define the number of clusters, and it uses tree-based illustrations to represent cluster results, called the dendrogram of the members.

K-means clusters groups into K groups. “K” represents the number of clusters the user must define the “K” in this method. Similar members form the same cluster; dissimilar members create dissimilar clusters. The main idea behind forming clusters is to minimize the total variation within a cluster. Each cluster is represented by its centroid, corresponding to the mean number of members. The initial K-means clustering step involves setting a specific number of clusters, known as K. The second step involves randomly selecting K members to serve as the initial centers for the clusters. These members are called cluster means or centroids. In the third step, each member is assigned to the nearest centroids according to the Euclidean distance between the centroids. Step four: Calculating new cluster centers: The new cluster centroids are

calculated using all the members. Steps 3 and 4 repeat iteratively to minimize the total within the sum of square distance till the assignment to a cluster stop changes or the defined number of iterations is met [43].

Two methods, agglomerative and divisive, exist in hierarchical clustering. Agglomerative clustering starts with each observation as its cluster, while divisive clustering works the opposite way, with each observation initially grouped into a single cluster. In agglomerative clustering, each observation is considered a separate cluster in the first place. Two clusters with similar structures are combined into a single cluster. The process is repeated until all observation values are brought together in a single cluster until all observation values form a single cluster. Unlike agglomerative clustering, analysis in divisive clustering, all observation values are initially considered as a single cluster in one. Observation values with different structures are divided into two clusters, forming two new clusters. The division continues until each observation value is represented by its cluster [44].

In this study, agglomerative hierarchical clustering methods were used. Euclidean or Manhattan distance is commonly employed to quantify the similarity or dissimilarity of observations. This study applied Euclidean distance to assess the dissimilarity between each pair of observations. Different linkage methods include *complete*, *single*, *average*, *centroid*, and *ward*.

Two clustering methods, K-means, and hierarchical clustering, were coded in R Programming language by using the help of stats [45], cluster [46], factoextra [47], gridextra [48], circlize [49], and ggplot2 [50] R Packages.

2.3.3. TOPSIS and MOORA MCDM methods

This section briefly explains the concept of MCDM, its purpose, and application areas, followed by a detailed discussion of the TOPSIS and MOORA methods. The MCDM process involves considering multiple criteria when selecting or ranking the best alternatives. Due to its wide range of tools and methods can be used in various fields, such as healthcare, finance, and engineering. Therefore, MCDM is essential for decision-makers to evaluate and select the best alternatives or rank them. MOORA and TOPSIS methods, as the most preferred MCDM method in literature, were chosen in this study to organize and compare food sweeteners according to their hazardous effects on cell DNA [51,52].

The MOORA method declared by Brauers and Zavadskas in 2006 (Multi-Objective Optimization Using Ratio Analysis) [53] is applied when the information for each criterion (length, intensity, and moment) for alternatives is obtained and combined to form a decision matrix. A sample decision matrix (*D*) for this problem with *n* criteria and *m* alternatives is shown below in Equation (1)

$$\boldsymbol{D} = \begin{bmatrix} \boldsymbol{d_{11}} & \boldsymbol{d_{12}} & \cdots & \boldsymbol{d_{1j}} & \cdots & \boldsymbol{d_{1n}} \\ \boldsymbol{d_{21}} & \boldsymbol{d_{22}} & \cdots & \boldsymbol{d_{2j}} & \cdots & \boldsymbol{d_{2n}} \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ \boldsymbol{d_{i1}} & \boldsymbol{d_{i2}} & \cdots & \boldsymbol{d_{ij}} & \cdots & \boldsymbol{d_{in}} \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ \boldsymbol{d_{m1}} & \boldsymbol{d_{m2}} & \cdots & \boldsymbol{d_{mj}} & \cdots & \boldsymbol{d_{mn}} \end{bmatrix}_{\boldsymbol{m \times n}} . \tag{1}$$

The criteria weights are denoted as *W* vectors with a sum of 1, as given in Equation (2) below

$$\boldsymbol{W} = [\boldsymbol{w_1} \;\; \boldsymbol{w_2} \;\; \cdots \;\; \boldsymbol{w_i} \;\; \cdots \;\; \boldsymbol{w_n}]_{\boldsymbol{1 \times n}} . \tag{2}$$

To normalize the decision matrix above, Equation (3) was used. $\boldsymbol{F} = \left[\boldsymbol{f_{ij}}\right]_{\boldsymbol{m \times n}}$, stands for normalized decision matrix

$$\boldsymbol{f_{ij}} = \frac{\boldsymbol{d_{ij}}}{\sqrt{\sum_{i=1}^{m} \boldsymbol{d_{ij}^2}}} . \tag{3}$$

All values of the normalized decision matrix are weighted by multiplying them by the importance weight of each criterion. $\boldsymbol{W} = \left[\boldsymbol{w_j}\right]_{\boldsymbol{1\times n}}$ is the weight vector of the criteria, and $\boldsymbol{R} = \left[\boldsymbol{r_{ij}}\right]_{\boldsymbol{m\times n}}$ is the weighted decision matrix; the formula used to calculate the value of; r_{ij} is given in Equation (4) below. In this study, each criterion (length, intensity, and moment) was weighted equally at 33.333%

$$\boldsymbol{r_{ij} = f_{ij} w_j}\,. \tag{4}$$

A reference point (r) was obtained from the values in the decision matrix of the problem to determine the reference point. This reference point is the solution derived from the best results for each criterion not included in the issue. When performing solutions according to the MOORA method, the values of the alternatives are evaluated based on their differences with this reference point. Equation (5) below is given for the case where all criteria are positive

$$\left\{\boldsymbol{r_j \mid r_j = \max_i\left(r_{ij}\right) \quad for \quad \forall j}\right\}. \tag{5}$$

d_i, was used to calculate relative values for the reference point shown in Equation (6) below to represent the maximum difference of each alternative concerning the reference point for the final rankings

$$\boldsymbol{d_i = \max_j\left(r_j - r_{ij}\right)}\,. \tag{6}$$

In the final rankings to be carried out after the maximum difference values are obtained for each alternative, the alternative with the lowest maximum difference is determined as the most ideal alternative. The alternative with the most considerable maximum difference from the reference point is identified as the last alternative to be preferred. For all rankings, the final difference values of the alternatives are ranked from most significant to most minor.

In 1981, the TOPSIS method was presented by Hwang and Yoon [54]. As part of this method, alternatives are ranked according to their distance from ideal and non-ideal solutions. The theory behind the technique states that the farther an alternative is from the non-ideal solution, the closer it is to the perfect solution. The distances between two points are calculated using Euclidean distance calculations.

The first four steps are the same as those used in the MOORA method. In the fifth stage, one ideal solution (A^+) and one non-ideal solution (A^-) are obtained from the values in the decision matrix of the problem. These solutions are derived from the best and worst results for each criterion, which are not included in the issue. The TOPSIS method evaluates alternatives based on their distance from these values. Equations (7) and (8) below are given for the case where all criteria are positive

$$\boldsymbol{A^+ = \left\{a_j^+ \mid a_j^+ = \max_i\left(r_{ij}\right) \quad for \quad \forall j\right\}} \tag{7}$$

$$\boldsymbol{A^- = \left\{a_j^- \mid a_j^- = \min_i\left(r_{ij}\right) \quad for \quad \forall j\right\}}\,. \tag{8}$$

Each alternative's distance to the ideal and non-ideal solutions will be calculated in this stage. In the TOPSIS method, utilizing Euclidean distances for distance calculation, S_i^+ is the distance of criterion i to the ideal solution, and S_i^- is the distance of criterion i to the non-ideal solution, as shown in Equations (9) and (10), respectively. Since one result is produced for each alternative, the sets S_i^+ and S_i^- will be vectors of size m alternatives

$$\boldsymbol{S_i^+ = \sqrt{\sum_{j=1}^{n}\left(r_{ij} - a_j^+\right)^2}} \tag{9}$$

$$\boldsymbol{S_i^- = \sqrt{\sum_{j=1}^{n}\left(r_{ij} - a_j^-\right)^2}}\,. \tag{10}$$

The method calculates the relative closeness values (C_i^+) of the alternatives for the final evaluations as follows in Equation (11), based on the logic that the alternative that is furthest from the non-ideal solution is the alternative that is closer to the ideal solution

$$C_i^+ = \frac{s_i^-}{s_i^+ + s_i^-}. \tag{11}$$

Since the relative closeness value for each alternative is calculated according to its distance from the non-ideal solution, the alternative with the most significant relative closeness value is the ideal alternative. All alternatives are ranked in descending order to evaluate the other alternatives in the problem. The alternative with the smallest C_i^+ value can be characterized as the last alternative to be preferred since it is the closest to the non-ideal solution.

Two different algorithms were developed in MATLAB to apply the TOPSIS and MOORA methods [55].

3. ANALYSIS AND FINDINGS

The data for each tail length, intensity, and moment were analyzed and summarized in Table 2. Table 3 provides descriptive statistics for tail length, intensity, and moment. The mean values show that the average tail length is 39.51, the average intensity is 9.18, and the average moment is 2.1, with considerable variation observed across the dataset. The maximum values suggest that certain sweeteners have significantly higher effects when compared to other ones. Higher correlations among all these comet parameters are given in Table 4. The length and moment are moderately correlated (38.4%), while length and intensity (19.6%), as well as intensity and moment (27.1%), show weaker correlations. This suggests that while these metrics are somewhat related, they capture different aspects of DNA damage. Lastly, Table 5 shows the correlations after excluding negative and positive control groups. According to Table 5, the correlations increased substantially, particularly between tail intensity and moment, with a correlation of 96.5%. In addition to that, the tail length and moment correlation was 68.4%. This suggests that sweeteners exhibit a more consistent relationship between these variables when negative and positive control groups are excluded. Our findings show that distinct characteristics of tail length, intensity, and moment influence overall correlation patterns.

Table 2. *Comet tail length, tail intensity, and tail moment averages*

sweeteners	length	intensity	moment
DSX_C	17.960	3.132	0.598
DSX_C+	43.090	15.324	3.172
DSX_M1	18.345	3.947	0.816
DSX_M2	18.615	4.071	0.956
DSX_M3	19.245	5.537	1.268
DSX_M4	24.915	9.075	1.972
DSX_M5	33.400	10.188	1.997
DSX_M6	34.520	10.668	2.176
DSX_M7	38.610	12.067	2.381
DSX_M8	40.040	13.315	2.973
DS_1000	40.515	10.841	2.617
DS_125	26.250	5.389	0.957
DS_250	27.680	8.901	1.854
DS_500	27.820	12.067	2.381
DS_62.5	24.915	4.500	0.914
DS_C	19.245	4.071	0.598
DS_C+	42.425	14.636	3.092
XYL_1000	60.010	11.983	2.379
XYL_125	53.535	8.440	1.872
XYL_250	53.905	9.308	2.212

Table 3. *Descriptive statistics for averages of tail length, tail intensity, and tail moment*

	length	intensity	moment
Min.	17.960	3.132	0.598
1st Qu.	24.910	5.463	1.112
Median	34.520	9.308	1.997
Mean	39.510	9.179	2.099
3rd Qu.	48.310	12.025	2.499
Max	132.390	15.323	6.656

Table 4. *Correlations between tail length, tail intensity, and tail moment*

	length	**intensity**	**moment**
length	1	19.6%	38.4%
intensity	19.6%	1	27.1%
moment	38.4%	27.1%	1

Table 5. *Correlations when DS_C, DSX_C, XYL_C, DS_C+, DSX_C+, and XYL_C+ excluded*

	length	**intensity**	**moment**
length	1	62.7%	68.4%
intensity	62.7%	1	96.5%

XYL_500	55.880	10.932	2.725
XYL_C	55.310	8.017	1.708
XYL_C+	132.390	14.711	6.656

moment	68.4%	96.5%	1

The histogram of comet assay data is given in Figure 2. In this figure, the distributions of the DNA damage values for three features are right-skewed. Most data accumulated at lower values, while larger values are less frequent.

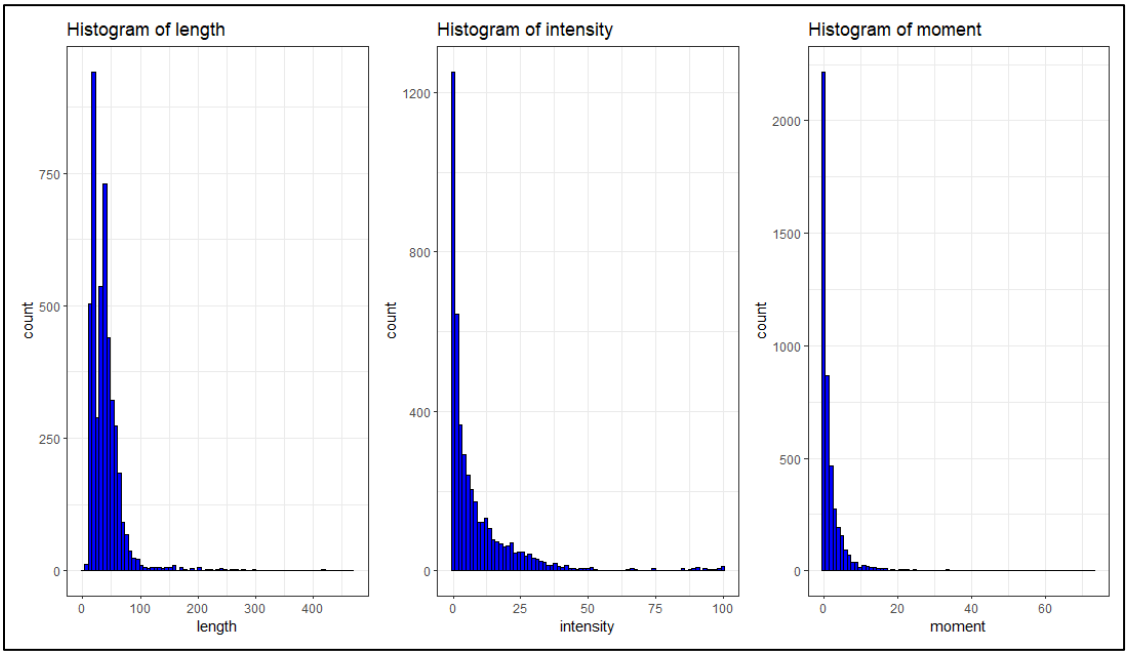


Figure 2. *Histogram of comet tail length, tail intensity, and tail moment*

Table 2 was visually represented in Figure 3 for comet tail length, tail intensity, and tail moment according to their average values. In addition to these results, line graphs stand for negative and positive control groups and individual and mixed sweeteners, according to their magnitudes on cell DNA damage. In Figure 3, DS_500, DS_1000, DSX_M7, and DSX_M8 µg/mL exhibited more extensive DNA damage than other sweetener concentrations for comet tail intensity. Additionally, DS_500, DS_1000, DSX_M8, and XYL_500 µg/mL showed more significant DNA damage over the others for comet tail moment.

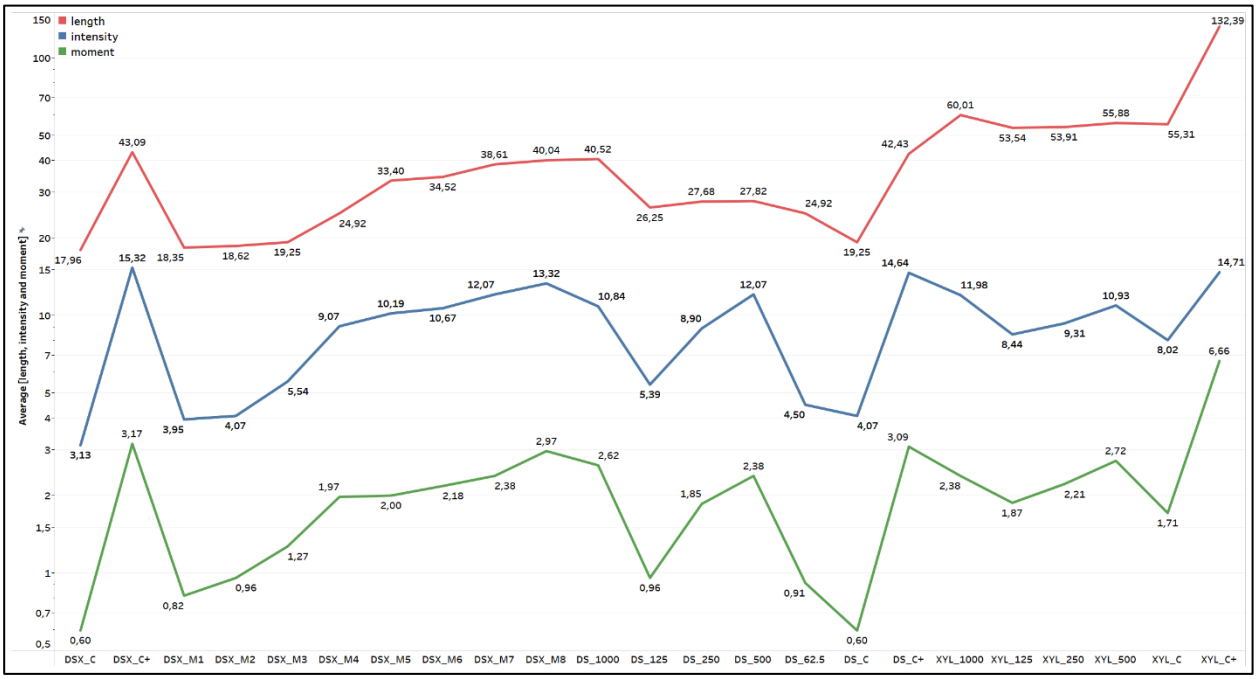


Figure 3. *Comet tail length, tail intensity, and tail moment line graph for sweeteners and mixed, including negative and positive control groups*

Figure 4 presents the separation references for sweeteners. Among the 23 applications, 10 exceeded the average tail length of 39.51, 12 surpassed the average tail intensity of 9.18, and 11 were above the average tail moment of 2.1. The features that exceed the average in length, intensity, and moment were identified separately. Among these features, the following were found to be shared across all three categories: DSX_C+, DSX_M8, DS_1000, DS_C+, XYL_1000, XYL_250, XYL_500, and XYL_C+. Additionally, some features were found to be unique to specific metrics. XYL_C and XYL_125 appeared exclusively in the length category, while DSX_M5 was unique to intensity. Notably, no features were uniquely above average, indicating a substantial overlap between the moment and other categories.

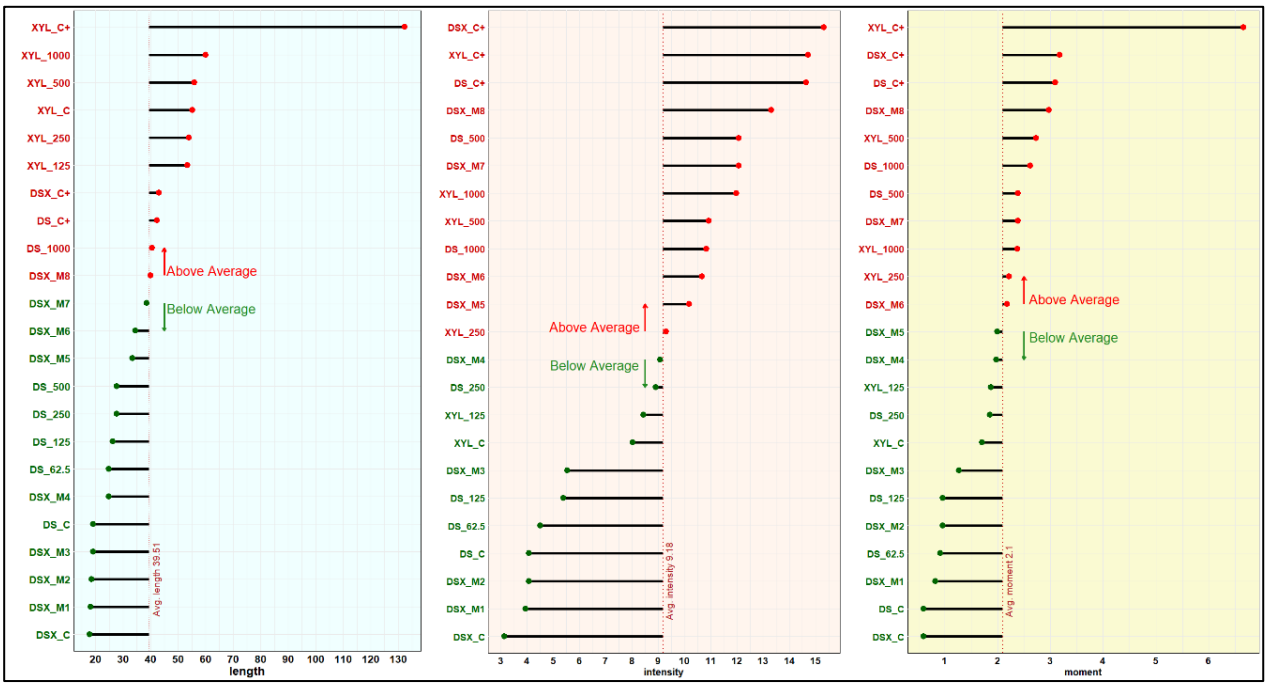


Figure 4. *Below or above averages for sweeteners/mixed and negative and positive control groups*

When the Kruskal-Wallis test was applied, the test results showed statistically significant differences among the groups, rejecting the null hypothesis (Table 6). After a significant difference between groups was revealed, a One-Way ANOVA test was performed thereafter.

Table 6. *Kruskal-Wallis test results*

Test Statistics[a,b]

	length	intensity	moment
Kruskal-Wallis H	2960.511	455.787	670.883
df	22	22	22
Asymp. Sig.	.000	.000	.000

a. Kruskal Wallis Test

b. Grouping Variable: content

One-way ANOVA was conducted to further analyze these differences, revealing significant variation in mean scores for tail length, tail intensity, and tail moment across food additive groups (Table 7).

Table 7. *One-way ANOVA test results*

ANOVA

		Sum of Squares	df	Mean Square	F	Sig.
length	Between Groups	2620577.205	22	119117.146	219.682	.000
	Within Groups	2481764.670	4577	542.225		
	Total	5102341.875	4599			
intensity	Between Groups	63306.270	22	2877.558	13.935	.000
	Within Groups	945148.055	4577	206.499		
	Total	1008454.325	4599			
moment	Between Groups	7148.377	22	324.926	14.972	.000
	Within Groups	99329.686	4577	21.702		
	Total	106478.063	4599			

Pairwise comparisons were performed to compare treatment groups with control groups, and the results were summarized in multiple comparison tables (Table 8), with significant differences highlighted in bold.

Table 8. *Comparison table for control groups vs. sweeteners/mixed*

length			intensity			moment		
DS_C	**XYL_C**	**DSX_C**	**DS_C**	**XYL_C**	**DSX_C**	**DS_C**	**XYL_C**	**DSX_C**
DSX_C	***DSX_C****	DSX_M1	DSX_C	***DSX_C****	DSX_M1	DSX_C	***DSX_C****	DSX_M1
DSX_M1	***DSX_M1****	DSX_M2	DSX_M1	***DSX_M1****	DSX_M2	DSX_M1	DSX_M1	DSX_M2
DSX_M2	***DSX_M2****	DSX_M3	DSX_M2	***DSX_M2****	DSX_M3	DSX_M2	DSX_M2	DSX_M3
DSX_M3	***DSX_M3****	***DSX_M4****	DSX_M3	DSX_M3	***DSX_M4****	DSX_M3	DSX_M3	DSX_M4
DSX_M4*	***DSX_M4****	***DSX_M5****	***DSX_M4****	DSX_M4	***DSX_M5****	DSX_M4	DSX_M4	DSX_M5
DSX_M5*	***DSX_M5****	***DSX_M6****	***DSX_M5****	DSX_M5	***DSX_M6****	***DSX_M5****	DSX_M5	***DSX_M6****
DSX_M6*	***DSX_M6****	***DSX_M7****	***DSX_M6****	DSX_M6	***DSX_M7****	***DSX_M6****	DSX_M6	***DSX_M7****
DSX_M7*	***DSX_M7****	***DSX_M8****	***DSX_M7****	DSX_M7	***DSX_M8****	***DSX_M7****	DSX_M7	***DSX_M8****
DSX_M8*	***DSX_M8****	***DSX_C**+*	***DSX_M8****	DSX_M8	***DSX_C**+*	***DSX_M8****	DSX_M8	***DSX_C**+*
***DSX_C**+*	***DSX_C**+*	DS_C	***DSX_C**+*	***DSX_C**+*	DS_C	***DSX_C**+*	DSX_C+	DS_C
DS_62.5*	***DS_C****	***DS_62.5****	DS_62.5	***DS_C****	DS_62.5	DS_62.5	***DS_C****	DS_62.5
DS_125*	***DS_62.5****	***DS_125****	DS_125	***DS_62.5****	***DS_125****	DS_125	DS_62.5	DS_125
DS_250*	***DS_125****	***DS_250****	***DS_250****	DS_125	***DS_250****	***DS_250****	***DS_125****	***DS_250****
DS_500*	***DS_250****	***DS_500****	***DS_500****	DS_250	***DS_500****	***DS_500****	DS_250	***DS_500****
DS_1000*	***DS_500****	***DS_1000****	***DS_1000****	DS_500	***DS_1000****	***DS_1000****	DS_500	***DS_1000****
***DS_C**+*	***DS_1000****	***DS_C**+*	***DS_C**+*	DS_1000	***DS_C**+*	***DS_C**+*	DS_1000	***DS_C**+*
XYL_C*	***DS_C**+*	***XYL_C****	***XYL_C****	DS_C+	***XYL_C****	***XYL_C****	***DS_C**+*	***XYL_C****
XYL_125*	XYL_125	***XYL_125****	***XYL_125****	XYL_125	***XYL_125****	***XYL_125****	XYL_125	***XYL_125****
XYL_250*	XYL_250	***XYL_250****	***XYL_250****	XYL_250	***XYL_250****	***XYL_250****	XYL_250	***XYL_250****
XYL_500*	XYL_500	***XYL_500****	***XYL_500****	XYL_500	***XYL_500****	***XYL_500****	XYL_500	***XYL_500****
XYL_1000*	XYL_1000	***XYL_1000****	***XYL_1000****	XYL_1000	***XYL_1000****	***XYL_1000****	XYL_1000	***XYL_1000****
***XYL_C**+*	***XYL_C**+*	***XYL_C**+*	***XYL_C**+*	***XYL_C**+*	***XYL_C**+*	***XYL_C**+*	***XYL_C**+*	***XYL_C**+*

Following the statistical analysis, clustering methods were applied to further explore patterns and relationships among the food sweeteners based on their impact on DNA damage. Clustering methods mainly depend on the similarities or dissimilarities between the observations. Heating maps were given in Figure 5 for four scenarios. Although it is apparent that the heating map for excluding negative and positive control groups is much more appropriate for clustering, we aimed to cluster sweeteners with negative and positive control groups included. The dark level of the blue states the dissimilarities and soft colors signify fair similarities. DSX_M1, DSX_M2, DSX_M3, DS_62.5, DS_125, DS_C, and DSX_C are similar in creating soft colors (red) when negative and positive control groups were included (Figure 5.a). When negative control was excluded, DSX_M1, DSX_M2, DSX_M3, DS_62.5, DS_125, and DS_C^{+} were similar (Figure 5.b). DSX_M1, DSX_M2, DSX_M3, DS_62.5, DS_125, DS_C, and DSX_C were similar when the positive control was excluded (Figure 5.c). When negative and positive control groups were excluded, DSX_M4, DSX_M5, DSX_M6, DS_250, and DS_500 were similar in creating soft colors (red). Moreover, DSX_M1, DSX_M2, DSX_M3, DS_65.5 and DS_125 were similar. DSX_M7, DSX_M8, DS_1000, XYL_125, XYL_250, XYL_500, and XYL_1000 were similar when negative and positive control groups were excluded (Figure 5.d).

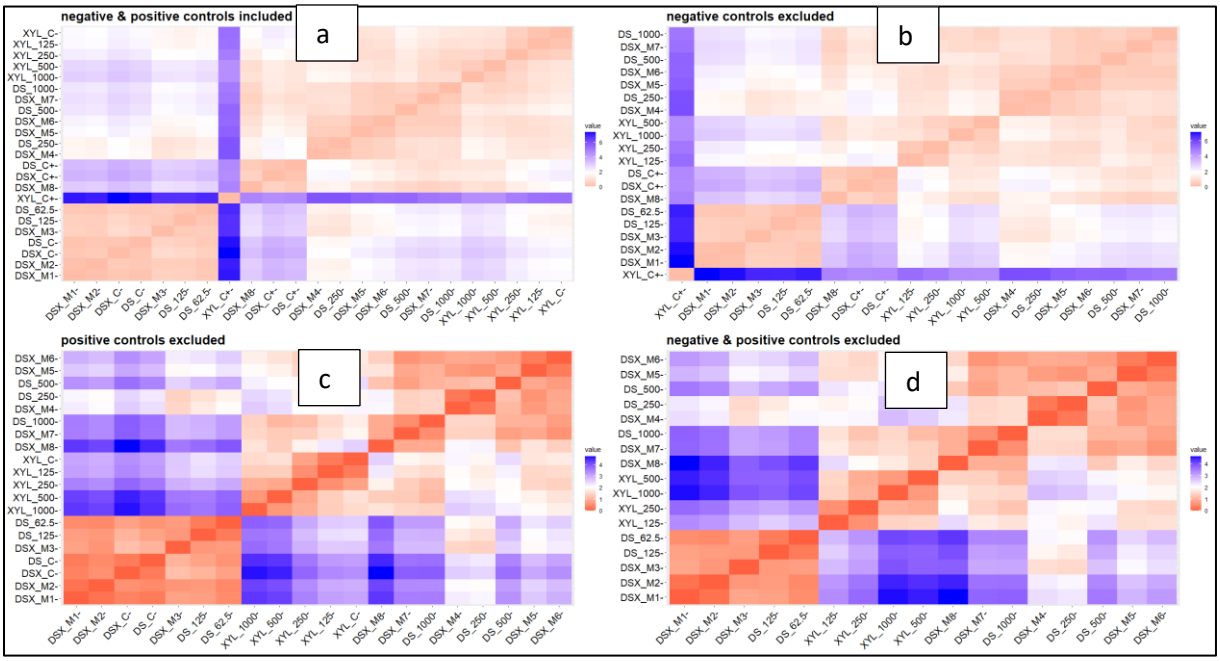


Figure 5. *Heating maps for groups*

Although to determine the optimal number of clusters (K), such as the Elbow Method and Silhouette Score are commonly used, all clustering visualizations from K=2 to 7 have been provided to offer a more detailed view of the clustering results. This approach allows readers to observe how sweeteners form the cluster with different values of K, facilitating a more comprehensive understanding of the clustering. Different *K* parameters were defined, and K-means clustering was performed to reveal forming clusters 2 to 7, which are given in Figures 6 and 7. XYL_C+ formed a discrete cluster for all the *K* parameters except for 2.

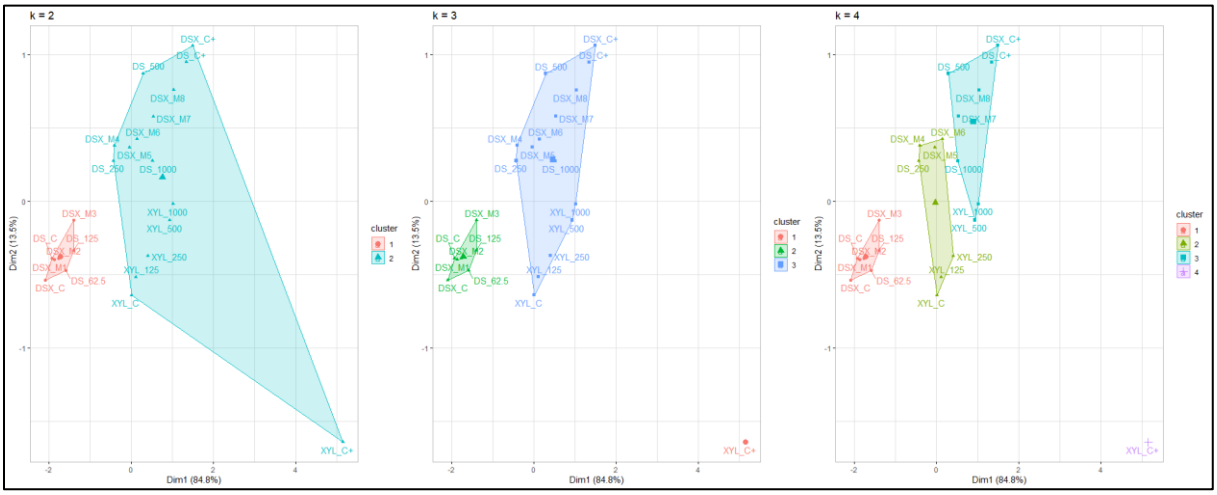


Figure 6. *Two to four clusters for groups*

When *K = 6* considered, DSX_C, DSX_M1, DSX_M2, DSX_M3, DS_125, DS_62.5, and DS_C formed cluster 1, XYL_500 and XYL_1000 formed cluster 2, DS_C+, DSX_C+, and DSX_M8 formed cluster 3, XYL_C, XYL_125, and XYL_250 formed cluster 4, DSX_M4, DSX_M5, DSX_M6, DSX_M7, DS_1000, DS_250, and DS_500 formed cluster 5. Finally, XYL_C+ formed cluster 6 since all the observations for comet tail length, tail intensity, and tail moment were higher than in other groups for XYL_C+. Since cluster numbers are expressed by a different label number each time the algorithm is run, one can expect that there will be different cluster number labels in the figures below (Figure 7). To sum up, similar observations form a cluster; on the other hand, dissimilar ones form a different cluster.

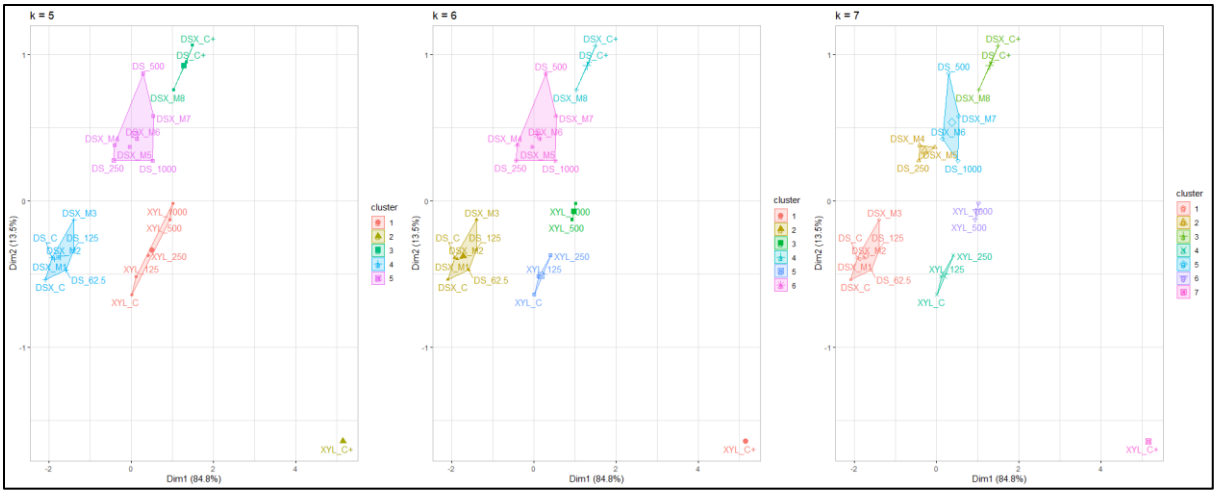


Figure 7. *Five to seven clusters for groups*

We calculated the agglomerative coefficients to discern the robustness of each clustering method discussed. Since values closer to 1 suggest robust clustering, average: 90.96%, single: 90.10%, complete: 91.82%, and ward: 92.67% were calculated, respectively. Since the most substantial cluster structure was obtained using the Ward method, a dendrogram plot was created using Ward's process, as shown in Figure 8.

In Figure 8, each leaf corresponds to one sweetener/mixed or control group in the circle dendrogram below. Observations that are alike each other are combined into branches. As the height increases, the similarity decreases. For instance, when the circle dendrogram (a) is considered, DSX_M1 and DSX_M2 are fused first, indicating that these observations are similar. As with other similar observations, DS_62.5 and DS_125 are fused secondly. In this circle dendrogram, 6 clusters were revealed by different colors in Figure 8.b.

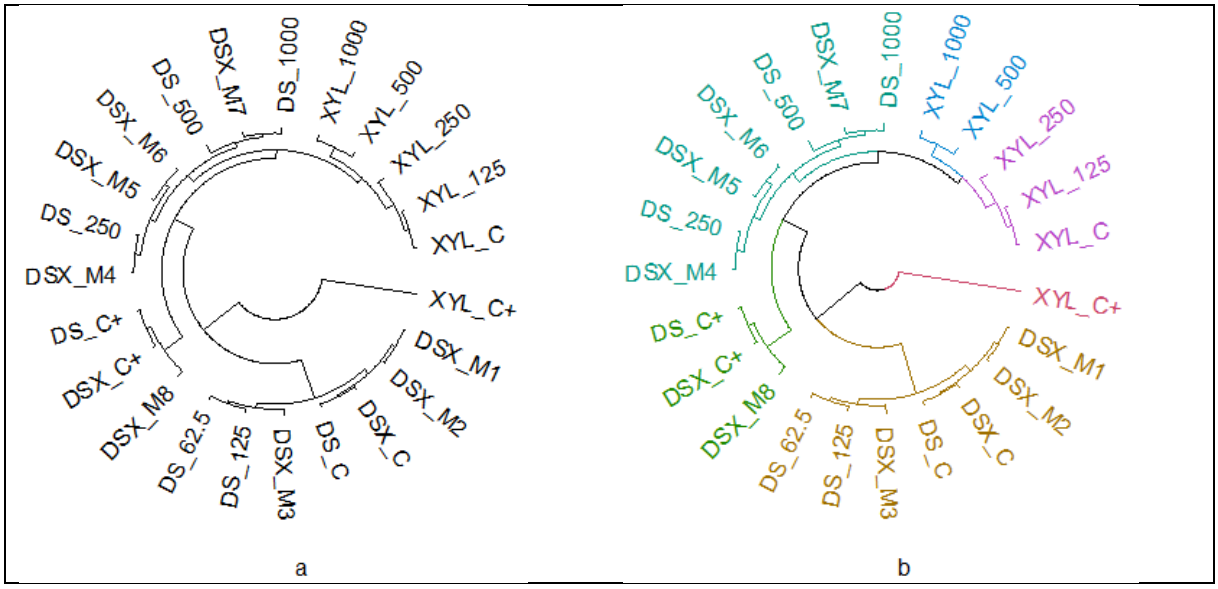


Figure 8. *Circle dendrogram of groups according to hierarchical clustering*

Cutting the dendrogram established by a hierarchical clustering algorithm at different levels allows different clustered results without reusing the hierarchical clustering algorithm. When hierarchical clustering was created according to six groups, it was similar to K-means results. Cluster 1 consists of DSX_C, DSX_M1, DSX_M2, DSX_M3, DS_125, DS_62.5, DS_C. Cluster 2 consists of XYL_500 and XYL_1000. Cluster 3 consists of DS_C+, DSX_C+, and DSX_M8. Cluster 4 consists of XYL_C, XYL_125, and XYL_250. Cluster 5 consists of DSX_M4, DSX_M5, DSX_M6, DSX_M7, DS_1000, DS_250, and DS_500. Cluster 6 consists of XYL_C+. Based on both clustering results, the mixture and DS concentrations form a larger group and cause DNA damage at closer levels. XYL concentrations form different groups. Therefore, the level of damage caused by XYL concentrations, especially at XYL_500 and XYL_1000 on DNA damage, differs from others (Figure 8). In Figure 9, K-means and hierarchical clustering resulted in the same clustering results in a more precise visual.

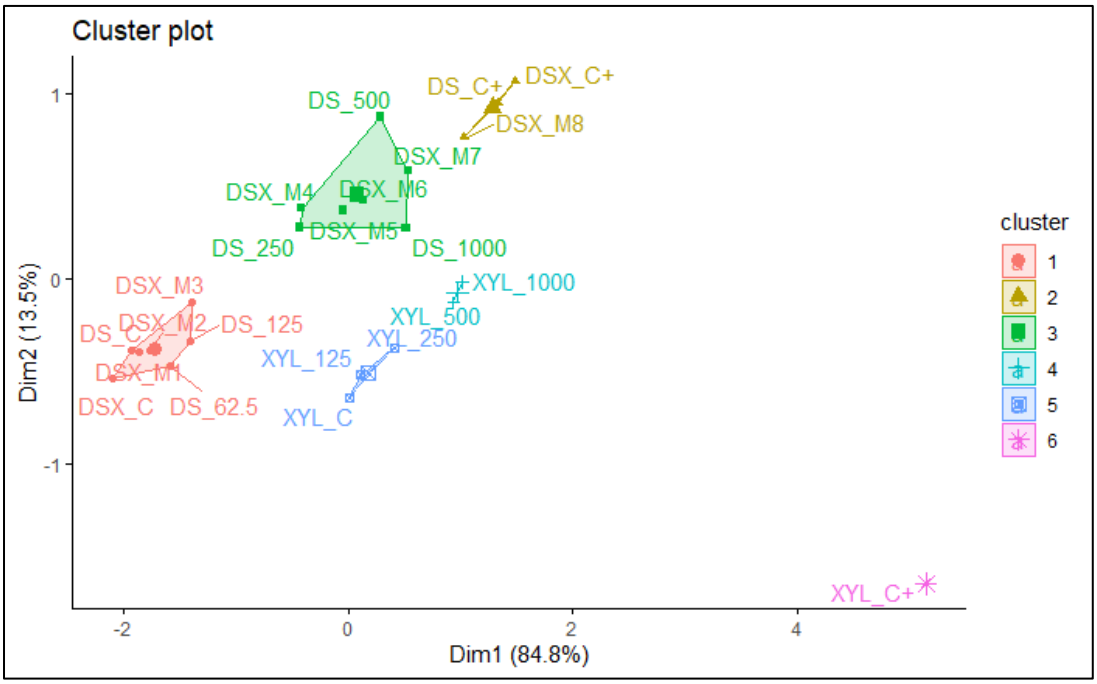


Figure 9. *Clustering result for K-means and hierarchical clustering*

After identifying clusters of food sweeteners based on their impact on DNA damage, MCDM methods were employed to further evaluate and rank them according to their hazardous effects. In Table 9, sweeteners,

negative control groups, and positive control groups were ranked with MOORA and TOPSIS techniques by their hazardous effects in terms of tail length, tail intensity, and tail moment. Higher xylitol concentrations (XYL 1000 µg/mL and XYL 500 µg/mL) were more genotoxic than DS and DSX at both methods. It is also shown that XYL_1000 generated more significant DNA damage when applied alone than when used in a mixture with sorbitol (DSX_M8). In contrast, DS_1000 caused less genotoxic effect than when combined with xylitol (DSX_M8), and according to the result of this study, XYL_1000, XYL_500, DSX_M8, and DS_1000 had more toxic effects than others, respectively.

Table 9. *Ranking results for sweeteners and control groups with MOORA and TOPSIS*

Ranking	MOORA		TOPSIS	
1	XYL_C+	1.476510	XYL_C+	0.983557
2	DSX_C+	0.788386	DSX_C+	0.415689
3	DS_C+	0.764045	DS_C+	0.401685
4	XYL_1000	0.726683	XYL_1000	0.382296
5	XYL_500	0.715463	XYL_500	0.378566
6	DSX_M8	0.715258	DSX_M8	0.37281
7	DS_1000	0.634929	DS_1000	0.32004
8	DSX_M7	0.631936	XYL_250	0.318267
9	XYL_250	0.628521	DSX_M7	0.315253
10	DS_500	0.58319	DS_500	0.290557
11	XYL_125	0.579469	XYL_125	0.285251
12	DSX_M6	0.566486	XYL_C	0.27818
13	XYL_C	0.564647	DSX_M6	0.27176
14	DSX_M5	0.536048	DSX_M5	0.250269
15	DSX_M4	0.472099	DSX_M4	0.213446
16	DS_250	0.47085	DS_250	0.207701
17	DS_125	0.313786	DSX_M3	0.096431
18	DSX_M3	0.311821	DS_125	0.0866872
19	DS_62.5	0.285389	DS_62.5	0.0649219
20	DSX_M2	0.251484	DSX_M2	0.0466786
21	DSX_M1	0.235688	DSX_M1	0.0324946
22	DS_C	0.223715	DS_C	0.0262153
23	DSX_C	0.198149	DSX_C	0

4. DISCUSSION

In the literature, no study utilizes ML algorithms and MCDM methods to cluster and rank the genotoxic effect of food sweeteners determined by the comet assay. In this study, we collected a dataset on comet assay to reveal and assess the genotoxic potential of two food sweeteners, DS and XYL, and their combination (DS+XYL=DSX) when they are used as a mix via ML and MCDM methods for the first time. ML is a subset of AI where algorithms are trained and developed to automate data-driven decision-making. This connection results from a multidisciplinary study between data scientists and healthcare professionals; the principal purpose is to accelerate the identification of groups at higher risk for decision-making [56].

One of the last points to be addressed in registering any new food/feed ingredient is genotoxicity, in which a positive result often leads to a compound being replaced by a different version or being terminated in the development process [57,58]. Therefore, the genotoxic effects of artificial sweeteners have been evaluated by the comet assay in various cells. In one of these studies, Fındıklı and Türkoğlu [59] examined the genotoxic potential of DS at 1.25, 2.5, and 5.0 ppm concentrations in human lymphocytes. They observed that the tail DNA and tail moment significantly increased at 2.5 ppm. The DNA-damaging effect of maternal sorbitol consumption in lactating Wistar rat offspring after 14 days of mother-to-offspring exposure was determined by Cardoso et al. [8]. *In vivo*, the comet assay revealed that pups exposed to DS showed a dose-dependent genotoxic response. While 0.15 mg/kg/day exhibited two-fold DNA damage over the control group, 1.5 mg/kg/day and 150.0 mg/kg/day presented up to four and fivefold damage than the control group [8]. In a study carried out by Metwaly et al. [60], rats were exposed to 200, 300, and 500

µg/mL DS to estimate the genotoxic effects using the comet assay, and it was determined that DS had no genotoxic effect on the rat lymphocytes compared to the control.

Humans may be exposed to not only one type of food sweetener but two or more types in their daily diets and beverages. When food sweeteners are consumed together, the cumulative effect of their mixtures may be modified via synergistic, additive, or antagonistic interactions among these components, which may alter their genetic impacts. De Oliveira et al. [61] implemented a study in which the genotoxic effects of sodium saccharin and sodium cyclamate sweeteners were examined alone and in combination using *A. cepa* chromosome aberration and micronucleus tests. Researchers used various concentrations of sodium saccharin, sodium cyclamate, and sodium saccharin + sodium cyclamate for 48, 72, and 168-hour exposure periods. Higher cytotoxic and mutagenic effects were reported following the combined saccharin and sodium cyclamate treatments compared to their single treatment. Therefore, sodium cyclamate and sodium saccharin may have synergistic effects when they are used as a mixture. Mukhopadhyay et al. [62] investigated the combined genotoxic potential of aspartame (ASPTM) and acesulfame-K (ASK) sweeteners using chromosomal aberrations in mouse bone marrow cells. According to test results, various mixtures of these two sweeteners are weakly clastogenic, but neither damaged cells nor chromosomal aberrations per cell were not significantly differed. These studies reveal that it is essential to determine the combined genotoxic effects of chemicals on human health. It is therefore recommended that the use of sweeteners, both alone and in combination, should be evaluated with further testing systems to understand their mechanisms of action.

In our previous study, four various concentrations of XYL (125, 250, 500, and 1000 µg/mL) were also examined in human lymphocytes *in vitro* by different genotoxicity assays, including comet [4]. The results revealed that XYL at high concentrations may be mutagenic. Concentrations of 500 and 1000 µg/mL significantly increased DNA damage in isolated lymphocytes compared to the negative control. According to the data previously obtained by Gunes [40], DS significantly increased DNA damage between 62.50 µg/mL and 1000 µg/mL for comet tail length and tail intensity, and at 250 µg/mL, 500 µg/mL and 1000 µg/mL for tail moment. DNA damage also significantly increased at the combined treatment of DS+XYL at DSX_M4, DSX_M5, DSX_M6, DSX_M7, and DSX_M8 in terms of all three comet parameters [40]. The data obtained from Gunes [40] and Avuloglu-Yilmaz et al. [4] were used for statistical analysis, ML-based clustering, and MCDM ranking methods for the recent study. These findings suggest that computational approaches can provide a clear insight into genotoxicity assessment by combining all the parameters determined in the comet assay. XYL_1000 consistently emerged as the most hazardous across all methodologies. The combination of DS and XYL (DSX_M8) exhibited increased toxicity compared to individual sweeteners. ML and MCDM methods proved to be highly effective in distinguishing genotoxic effects.

The mean values (Table 2) of previously examined data used to make descriptive statistics for averages (Table 3) and correlations among all comet parameters (Table 4) were the initial steps of this study. When all the treatment and control groups were compared, the most correlation was found between tail length and tail moment. On the other hand, if control groups were excluded, the most correlated parameters were tail intensity and tail moment since the control groups mostly led to an increase in the formation of DNA migration in the tail length, while the sweetener treatment groups caused a rise in tail intensity. The difference might result from the positive control groups (H_2O_2) forming oxidative stress by generating hydroxyl-free radicals ($OH^{\bullet}$) that attack the DNA backbone and form single-strand breaks [63].

This study presented a statistically significant difference in the DNA damaging effect between the food sweeteners and control groups (Tables 6 and 7). Pairwise comparisons revealed the differences between the DS_C, XYL_C, and DSX_C groups and food sweeteners (Table 8). In addition to that, food sweeteners with similar effects have been identified using clustering methods. XYL_C+ formed a cluster by itself, which shows the dissimilarity from the other groups (Figures 6 and 7). Also, to rank food additives in terms of their effects on tail length, tail intensity, and tail moment, MOORA and TOPSIS methods were used together, and the three most DNA-damaging food sweetener treatments were identified as XYL_1000, XYL_500, and DSX_M8, respectively, when negative and positive control groups were excluded (Table 9). The results determined in this MCDM study coincided with the previous one, indicating that XYL_1000

and XYL_500 were ranked first and second highest in the DNA damage ranking of all three comet parameters. Similarly, another computational method, molecular docking, revealed that XYL had a high affinity for the DNA structure and crucial proteins (p53, ATM), corroborating their genotoxic potential [64].

MCDM methods provide crucial tools to rank and select the most appropriate among multiple alternatives [65]. Compared to other MCDM methods, the MOORA is known to have a high level of reliability [66,67]. The TOPSIS is another prevalent MCDM method [68]. No study using MCDM methods was found when this type of genotoxicity study was carried out. This study differs from genotoxicity studies in terms of the technique it uses. In this regard, this investigation will contribute to developing and increasing the existing literature. On the other hand, similar results in both methods show that experimental methods and multiple decision-making methods support each other.

In conclusion, this study is the first to assess previously determined comet values obtained in our laboratory to compare and rank various concentrations of two single food sweeteners and their mixture, and negative and positive control groups using MOORA and TOPSIS techniques. This novel and combined approach that integrates Kruskal Wallis and One-Way ANOVA methods, a novel combination of K-Means and Hierarchical Clustering, along with MOORA and TOPSIS, can predict the level and ranking of the toxic effect of chemicals in terms of DNA damage. This procedure allows for an innovative ranking of damaging effects based on comet assay data, including 13800 observations. Applying *in silico* methods in the risk assessment of food ingredients is rare in the food industry, possibly due to a need for more expertise in using these tools and analyzing results determined [69,70]. This study not only ranked and grouped the genotoxic effects of sweeteners, both alone and in combination forms, on human lymphocytes but also demonstrated that several novel statistical analyses, ML approaches, and MCDM methods could be applied more comprehensively in a more interpretable manner in genotoxicity studies rather than traditional statistical methods, which are believed to be an essential contribution to the literature in multidisciplinary studies. Integrating the Comet Assay with statistical methods (such as Kruskal-Wallis and One-Way ANOVA) and machine learning techniques (such as K-means clustering) can significantly enhance the robustness and interpretability of genotoxicity data. Advanced analytical approaches allow for more precise grouping and ranking of substances based on their genotoxic effects, providing valuable insights into their relative risks. This combined methodology may offer a more comprehensive and reliable way of assessing genotoxicity, contributing to more informed regulatory decisions and developing safer chemical compounds.

CONFLICTS OF INTEREST

No conflict of interest was declared by the authors.

REFERENCES

[1] Wu, L., Zhang, C., Long, Y., Chen, Q., Zhang, W., Liu, G., "Food additives: From functions to analytical methods", Critical Reviews in Food Science and Nutrition, 62(30): 8497-8517, (2022). DOI: https://doi.org/10.1080/10408398.2021.1929823

[2] Antonik, N., Janda, K., Jakubczyk, K., "Characteristics of sweeteners used in foods and their effects on human health", Pomeranian Journal of Life Sciences, 66(3): 57-65, (2020). DOI: https://doi.org/10.21164/pomjlifesci.723

[3] Chen, L., Zhang, Y., Zhou, Y., Shi, D., Feng, X.S., "Sweeteners in food samples: An update on pretreatment and analysis techniques since 2015", Food Chemistry, 408: 135248, (2023). DOI: https://doi.org/10.1016/j.foodchem.2022.135248

[4] Avuloglu-Yılmaz, E., Yuzbaşıoglu, D., Unal, F., "Investigating *in vitro* genotoxic effects of sweetener xylitol", KSU Journal of Agriculture and Nature, 25(6): 1315-1325, (2022). DOI: https://doi.org/10.18016/ksutarimdoga.vi.99382

[5] Yin, K.J., Xie, D.Y., Zhao, L., Fan, G., Ren, J.N., Zhang, L.L., Pan, S.Y., "Effects of different sweeteners on behavior and neurotransmitters release in mice", Journal of Food Science and Technology, 57: 113-121, (2020). DOI: https://doi.org/10.1007/s13197-019-04036-6

[6] Hernández-Pérez, A.F., Jofre, F.M., de Souza Queiroz, S., de Arruda, P.V., Chandel, A.K., de Almeida Felipe, M.D.G., Biotechnological Production of Bioactive Compounds, Verma, ML., Chandel, AK., Elsevier, 261-292, (2020).

[7] Rice, T., Zannini, E.K., Arendt, E., Coffey, A.A., "A review of polyols-biotechnological production, food applications, regulation, labeling, and health effects", Critical Reviews in Food Science and Nutrition, 60(12): 2034-2051, (2020). DOI: https://doi.org/10.1080/10408398.2019.1625859

[8] Cardoso, F.S., Araujo-Lima, C.F., Aiub, C.A., Felzenszwalb, I., "Exposure to sorbitol during lactation causes metabolic alterations and genotoxic effects in rat offspring", Toxicology Letters, 260: 36-45, (2016). DOI: https://doi.org/10.1016/j.toxlet.2016.08.018

[9] Roze, M., Crucean, D., Diler, G., Rannou, C., Catanéo, C., Jonchère, C., Le-Bail, A., Le-Bail, P., "Impact of maltitol and sorbitol on technological and sensory attributes of biscuits", Foods, 10(11): 2545, (2021). DOI: https://doi.org/10.3390/foods10112545

[10] Zhang, W., Chen, J., Chen, Q., Wu, H., Mu, W., "Sugar alcohols derived from lactose: lactitol, galactitol, and sorbitol", Applied Microbiology and Biotechnology, 104: 9487-9495, (2020). DOI: https://doi.org/10.1007/s00253-020-10929-w

[11] Awuchi, C.G., Echeta, K.C., "Current developments in sugar alcohols: Chemistry, nutrition, and health concerns of sorbitol, xylitol, glycerol, arabitol, inositol, maltitol, and lactitol", International Journal of Advanced Academic Research, 5(11): 1-33, (2019).

[12] Dasgupta, D., Ahuja, V., Singh, R., More, S., Mudliar, S., Kumar, M., "Food-grade xylitol production from corncob biomass with acute oral toxicity studies", World Journal of Microbiology and Biotechnology, 39(4): 102, (2023). DOI: https://doi.org/10.1007/s11274-023-03542-2

[13] Mäkinen, K.K., "Gastrointestinal disturbances associated with the consumption of sugar alcohols with special consideration of xylitol: scientific review and instructions for dentists and other healthcare professionals", International Journal of Dentistry, 2016(1): 967907, (2016). DOI: https://doi.org/10.1155/2016/5967907

[14] FDA., "Generally Recognized as Safe (GRAS)", U.S. Food & Drug Administration, (2017).

[15] Recoules, C., Touvier, M., Pierre, F., Audebert, M., "Evaluation of the toxic effects of food additives, alone or in mixture, in four human cell models", Food and Chemical Toxicology, 196: 115198, (2025). DOI: https://doi.org/10.1016/j.fct.2024.115198.

[16] Moriconi, E., Feraco, A., Marzolla, V., Infante, M., Lombardo, M., Fabbri, A., Caprio, M., "Neuroendocrine and metabolic effects of low-calorie and non-calorie sweeteners", Frontiers in Endocrinology, 11: 444, (2020). DOI: https://doi.org/10.3389/fendo.2020.00444

[17] Collins, A., Møller, P., Gajski, G., Vodenková, S., Abdulwahed, A., Anderson, D., Bankoglu, EE., Bonassi, S., Boutet-Robinet, E., Brunborg, G., Chao, C., Cooke, MS., Costa, C., Costa, S., Dhawan, A., de Lapuente, J., Del Bo', C., Dubus, J., Dusinska, M., Duthie, SJ., El Yamani, N., Engelward, B., Gaivão, I., Giovannelli, L., Godschalk, R., Guilherme, S., Gutzkow, KB., Habas, K., Hernández, A., Herrero, O., Isidori, M., Jha, AN., Knasmüller, S., Kooter, IM., Koppen, G., Kruszewski, M., Ladeira, C., Laffon, B., Larramendy, M., Le Hégarat, L., Lewies, A., Lewinska, A., Liwszyc, GE., López de Cerain, A., Manjanatha, M., Marcos, R., Milić, M., Moraes de Andrade, V., Moretti, M., Muruzabal, D., Novak, M., Oliveira, R., Olsen, AK., Owiti, N., Pacheco, M., Pandey, AK., Pfuhler,

S., Pourrut, B., Reisinger, K., Rojas, E., Rundén-Pran, E., Sanz-Serrano, J., Shaposhnikov, S., Sipinen, V., Smeets, K., Stopper, H., Teixeira, J. P., Valdiglesias, V., Valverde, M., van Acker, F., van Schooten, FJ., Vasquez, M., Wentzel, J. F., Wnuk, M., Wouters, A., Žegura, B., Zikmund, T., Langie, S. A. S., Azqueta, A., "Measuring DNA modifications with the comet assay: A compendium of protocols", Nature Protocols, 18(3): 929-989, (2023). DOI: https://doi.org/10.1038/s41596-022-00754-y

[18] Erikel, E., Yuzbasioglu, D., Unal, F., "Genotoxic and antigenotoxic potential of amygdalin on isolated human lymphocytes by the comet assay", Journal of Food Biochemistry, 44(10): e13436, (2020). DOI: https://doi.org/10.1111/jfbc.13436

[19] Žegura, B., Filipič, M., "Optimization in Drug Discovery", Yan, Z., Caldwell, GW., Humana Press, 301-313, (2004).

[20] Anderson, D., Dhawan, A., Laubenthal, J., "Genotoxicity Assessment. Methods in Molecular Biology", Dhawan A, Bajpayee M, Humana, 259-274, (2019).

[21] Azqueta, A., Ladeira, C., Giovannelli, L., Boutet-Robinet, E., Bonassi, S., Neri, M., Gajski, G., Duthie, S., Del Bo, C., Riso, P., Koppen, G., Basaran, N., Collins, A., Møller, P., "Application of the comet assay in human biomonitoring: An hCOMET perspective", Mutation Research/Reviews in Mutation Research, 78: 108288, (2020). DOI: https://doi.org/10.1016/j.mrrev.2019.108288

[22] Ceppi, M., Smolkova, B., Staruchova, M., Kazimirova, A., Barancokova, M., Volkovova, K., Collins, A., Kocan, A., Dzupinkova, Z., Horska, A., Buocikova, V., Tulinska, J., Liskova, A., Mikusova, M.L., Krivosikova, Z., Wsolova, L., Kuba, D., Rundén-Pran, E., El Yamani, N., Longhin., E., Dusinska, M., "Genotoxic effects of occupational exposure to glass fibres-A human biomonitoring study", Mutation Research/Genetic Toxicology and Environmental Mutagenesis, 885: 503572, (2023). DOI: https://doi.org/10.1016/j.mrgentox.2022.503572

[23] Marciano, L.P., Costa, L.F., Cardoso, N.S., Freire, J., Feltrim, F., Oliveira, G.S., Paula, F.B.A., Silvério, A.C.P., Martins, I., "Biomonitoring and risk assessment of human exposure to triazole fungicides", Regulatory Toxicology and Pharmacology, 147: 105565, (2024). DOI: https://doi.org/10.1016/j.yrtph.2024.105565

[24] Fan, D., Yang, H., Li, F., Sun, L., Di, P., Li, W., Tang, Y., Liu, G., "*In silico* prediction of chemical genotoxicity using machine learning methods and structural alerts", Toxicology Research, 7: 211-220, (2018). DOI: https://doi.org/10.1039/c7tx00259a

[25] Sizochenkoa, N., Syzochenko, M., Fjodorova, N., Rasulev, B., Leszczynski, J., "Evaluating genotoxicity of metal oxide nanoparticles: Application of advanced supervised and unsupervised machine learning techniques", Ecotoxicology and Environmental Safety, 185: 109733, (2019). DOI: https://doi.org/10.1016/j.ecoenv.2019.109733

[26] Xu, J.L., Lin, X., Gowen, A.A., "Combining machine learning with meta-analysis for predicting cytotoxicity of micro-and nanoplastics", Journal of Hazardous Materials Advances, 8: 100175, (2022). DOI: https://doi.org/10.1016/j.hazadv.2022.100175

[27] Ji, Z., Guo, W., Wood, E.L., Liu, J., Sakkiah, S., Xu, X., Patterson, T.A., Hong, H., "Machine Learning Models for Predicting Cytotoxicity of Nanomaterials", Chemical Research and Toxicology, 35: 125-139, (2022). DOI: https://doi.org/10.1021/acs.chemrestox.1c00310

[28] Atila, Ü., Baydilli, Y.Y., Sehirli, E., Turan, M.K., "Classification of DNA damages on segmented comet assay images using convolutional neural network", Computer Methods and Programs in Biomedicine, 186: 105192, (2020). DOI: https://doi.org/10.1016/j.cmpb.2019.105192

[29] Hong, Y., Han, H.J., Lee, H., Lee, D., Ko, J., Hong, Z.Y., Lee, J.Y., Seok, J.H., Lim, H.S., Son, W.C., Sohn, I., "Deep learning method for comet segmentation and comet assay image analysis", Scientific Reports, 10: 18915, (2020). DOI: https://doi.org/10.1038/s41598-020-75592-7

[30] Grenet, I., Yin, Y., Comet, J.P., Gelenbe, E., "Artificial neural networks and machine learning", 27th International Conference on Artificial Neural Networks, Rhodes, Greece, 335-345, (2018).

[31] Çelik, Ö., Altunaydın, S.S., "A research on machine learning methods and its applications", Journal of Educational Technology and Online Learning, 1(3): 25-40, (2018). DOI: https://doi.org/10.31681/jetol.457046

[32] FDA., "Artificial Intelligence and Machine Learning (AI/ML) for Drug Development" https://www.fda.gov/about-fda/center-drug-evaluation-and-research-cder/artificial-intelligence-drug-development. Access date: 08.08.2025

[33] Guo W., Liu J., Dong F., Song M., Li, Z., Khan, M.K.H., Patterson, T.A., Hong, H., "Review of machine learning and deep learning models for toxicity prediction", Experimental Biology and Medicine, 248: 1952-1973, (2023). DOI: https://doi.org/10.1177/15353702231209421

[34] Fan, D., Yang, H., Li, F., Sun, L., Di, P., Li, W., Tang, Y., Liu, G., "*In silico* prediction of chemical genotoxicity using machine learning methods and structural alerts", Toxicology Research, 7(2): 211-220, (2018). DOI: https://doi.org/10.1039/c7tx00259a

[35] Yadav, S., Soni, A., "Application of MOORA and TOPSIS for genotoxicity assessment: A decision-making approach", Journal of Hazardous Materials, 374: 245-253, (2019).

[36] Sanabria, P., Lumbaque, E.C., Becker, R.W., Jachstet, L.A., Scunderlick, D., Ruiz-Padillo, A., Sirtori, C., "Use of multi-criteria ranking method for environmental risk assessment of antineoplastic agents and their transformation products", Journal of Environmental Chemical Engineering, 11:109588, (2023). DOI: https://doi.org/10.1016/j.jece.2023.109588

[37] Singh, N.P., McCoy, M.T., Tice, R.R., Schneider, E.L., "A simple technique for quantitation of low levels of DNA damage in individual cells", Experimental Cell Research, 175: 184-191, (1988). DOI: https://doi.org/10.1016/0014-4827(88)90265-0

[38] Avuloglu Yilmaz, E., Yuzbasioglu, D., Unal, F., "Investigation of genotoxic effect of octyl gallate used as an antioxidant food additive in *in vitro* test systems", Mutagenesis, 38(3): 151-159, (2023). DOI: https://doi.org/10.1093/mutage/gead005

[39] Yuzbasioglu, D., Dilek, U.K., Erikel, E., Unal, F., "Antigenotoxic effect of hyperoside against Mitomycin C and hydrogen peroxide-induced genotoxic damage on human lymphocytes", Toxicology in Vitro, 90: 105604, (2023). DOI: https://doi.org/10.1016/j.tiv.2023.105604

[40] Gunes, B., "Gıda Tatlandırıcıları Olan D-Sorbitol ve D-Sorbitol + Ksilitol Karışımının İnsan Lenfositlerinde *in vitro* Genotoksik Etkileri", Yüksek Lisans Tezi (Turkish), Gazi Üniversitesi, Fen Bilimleri Enstitüsü, (2024).

[41] Unal, F., Saygılı, Y., Akbas, E., "Genetic Toxicology", F Unal, D Yuzbasioglu, Nobel Academic Publishing, (Turkish), 313-340, (2022).

[42] Akbas, E., Unal, F., Yuzbasioglu, D., "Cellular toxicities of gadolinium-based contrast agents used in magnetic resonance imaging", Journal of Applied Toxicology, 43: 958-972, (2023). DOI: https://doi.org/10.1002/jat.4416

[43] Özdemir, M., "R ile Programlama ve Makine Öğrenmesi", Nobel Akademik Yayıncılık, 280, (2020a).

[44] Özdemir, M., "R ile Programlama ve Makine Öğrenmesi", Nobel Akademik Yayıncılık, (2020b).

[45] Core Team R. *R:* A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, (2023).

[46] Maechler, M., Rousseeuw, P., Struyf, A., Hubert, M., Hornik, K., "Cluster: Cluster analysis basics and extensions. R package version 2.1.4. 2022. For new features, see the 'Changelog' file (in the package source)" https://CRAN.R-project.org/package=cluster

[47] Kassambara, A., Mundt, F., "Factoextra: Extract and visualize the results of multivariate data analyses. R Package Version 1.0.7." (2020). https://CRAN.R-project.org/package=factoextra

[48] Auguie, B., Antonov, A., "gridExtra: Miscellaneous functions for "Grid" Graphics. R Package Version 2.3." (2017). https://CRAN.R-project.org/package=gridExtra

[49] Gu, Z., Gu, L., Eils, R., Schlesner, M., Brors, B., "Circlize implements and enhances circular visualization in R. Bioinformatics", 30(19): 2811-2812, (2014). DOI: https://doi.org/10.1093/bioinformatics/btu393

[50] Wickham, H., "ggplot2: Elegant Graphics for Data Analysis", Springer-Verlag New York, (2016).

[51] Çelikbilek, Y., Özdemir, M., "Çok Kriterli Karar Verme Yöntemleri Açıklamalı ve Karşılaştırmalı Sağlık Bilimleri Uygulamaları", Nobel Akademik Yayıncılık, 337, (2020a).

[52] Çelikbilek, Y., Özdemir, M., "Çok Kriterli Karar Verme Yöntemleri Açıklamalı ve Karşılaştırmalı Sağlık Bilimleri Uygulamaları", Nobel Akademik Yayıncılık, 375, (2020b).

[53] Brauers, W.K., Zavadskas, E.K., "The MOORA method and its application to privatization in a transition economy", Control Cybern, 35(2): 445-469, (2006).

[54] Hwang, C.L., Yoon, K., "TOPSIS (technique for order preference by similarity to ideal solution)-a multiple attribute decision making, w: Multiple attribute decision making-methods and applications, a state-of-the-at survey", Berlin: Springer Verlag 128 (1981): 140.

[55] The MathWorks Inc. MATLAB version: 23.2.0.2489961 (R2023b), The MathWorks Inc. (2023). https://www.mathworks.com

[56] Rodríguez-Belenguer, P., Piñana, J.L., Sánchez-Montañés, M., Soria-Olivas, E., Martínez-Sober, M., Serrano-López, A.J., "A machine learning approach to identify groups of patients with hematological malignant disorders", Computer Methods and Programs in Biomedicine, 246: 108011, (2024). DOI: https://doi.org/10.1016/j.cmpb.2024.108011

[57] EFSA., "Scientific opinion on genotoxicity testing strategies applicable to food and feed safety assessment", EFSA Journal, 9(9): 2379, (2011).

[58] FDA., "Guidance for industry and other stakeholds toxicological principles for the safety assessment of food ingredients redbook", (2000). https://www.fda.gov/regulatory-information/search-fda-guidance-documents/guidance-industry-and-other-stakeholders-redbook-2000 link ekledim hocam

[59] Fındıklı, Z., Türkoğlu, Ş., "Determination of the effects of some artificial sweeteners on human peripheral lymphocytes using the comet assay", Journal of Toxicology and Environmental Health Sciences, 6(8): 147-153, (2014). DOI: https://doi.org/10.5897/JTEHS2014.0313

[60] Metwaly, A., Aboul-Enein, A., Abd-Allah, A., Hanafy, E., "Do synthetic food additives possess higher genotoxic effects than natural ones?", Bioscience Research, 15(4): 3329-3336, (2018).

[61] De Oliveira, V.A., Alves de Oliveira, V.M., de Oliveira, T.W.N., Damasceno, A.N.C., de Oliveira Silva, C.E., Medeiros, S.R.A., Soares, B.M., da Silva, F.C.C., de Aguiar, R.P.S., Islami M.T., Melo-Cavalcante, A.A.C., Peron, A.P., Castro e Sousa, J.M., "Evaluation of cytotoxic and mutagenic effects of two artificial sweeteners by using eukaryotic test systems", African Journal of Biotechnology, 16(11): 547-551, (2017). DOI: https://doi.org/10.5897/AJB

[62] Mukhopadhyay, M., Mukherjee, A., Chakrabarti, J., "*In vivo* cytogenetic studies on blends of aspartame and acesulfame-K", Food and Chemical Toxicology, 38(1): 75-77, (2000). DOI: https://doi.org/10.1016/S0278-6915(99)00115-5

[63] Benhusein, G., Mutch, E., Aburawi, S., Williams, F., "Genotoxic effect induced by hydrogen peroxide in human hepatoma cells using comet assay", Libyan Journal of Medicine, 5: 4637, (2020). DOI: https://doi.org/10.3402/ljm.v5i0.4637

[64] Okus, F., Yuzbasioglu, D., Unal, F., "Molecular docking study of frequently used food additives for selected targets depending on the chromosomal abnormalities they cause", Toxicology, 5: 153716, (2024). DOI: https://doi.org/10.1016/j.tox.2023.153716

[65] Çelikbilek, Y., Tüysüz, F., "An in-depth review of theory of the TOPSIS method: An experimental analysis", Journal of Management Analytics, 7(2): 281-300, (2020). DOI: https://doi.org/10.1080/23270012.2020.1748528

[66] Brauers, W.K.M., Zavadskas, E.K., "Robustness of MULTIMOORA: A method for multi-objective optimization", Informatica, 23(1) :1-25, (2012). DOI: https://doi.org/10.15388/Informatica.2012.346

[67] Sevim, F., Ugurluoglu Aldogan, E., "Evaluation of health systems performance of OECD countries using MOORA method", Journal of Health Management, 26(1): 172-183, (2024). DOI: https://doi.org/10.1177/09720634231215131

[68] Zyoud, S.H., Fuchs-Hanusch, D., "A bibliometric-based survey on AHP and TOPSIS techniques", Expert Systems with Applications, 78: 158-181, (2017). DOI: https://doi.org/10.1016/j.eswa.2017.02.016

[69] Blaauboer, B.J., Boobis, A.R., Bradford, B., Cockburn, A., Constable, A., Daneshian, M., Edwards, G., Garthoff, J.A., Jeffery, B., Krul, C., Schuermans, J., "Considering new methodologies in strategies for safety assessment of foods and food ingredients", Food and Chemical Toxicology, 91: 19-35, (2016). DOI: https://doi.org/10.1016/j.fct.2016.02.019

[70] Lo Piparo, E.L., Worth, A., Manibusan, M., Yang, C., Schilter, B., Mazzatorta, P., Jacobs, M.N., Steinkellner, H., Mohimont, L., "Use of computational tools in the field of food safety", Regulatory Toxicology and Pharmacology, 60(3): 354-362, (2011). DOI: https://doi.org/10.1016/j.yrtph.2011.05.003