PHARMACIA

Pharmacia 71: 1–10
DOI 10.3897/pharmacia.71.e127293

Review Article

Lavender essential oils–hidden relationships between the samples of origin

Ekaterina Kozuharova[1], Vasil Simeonov[2], Christina Stoycheva[1], Niko Benbassat[1], Daniela Batovska[3]

1 *Medical University of Sofia, Department of Pharmacognosy, Faculty of Pharmacy, Dunav 2 Sofia 1000, Bulgaria*
2 *University of Sofia "St. Kl. Okhridski", Faculty of Chemistry and Pharmacy, J. Boucher Blvd. Sofia 1164, Bulgaria*
3 *Institute of Chemical Engineering, Bulgarian Academy of Sciences, Acad. G. Bonchev Str., Bl. 103, 1113 Sofia, Bulgaria*

Corresponding author: Ekaterina Kozuharova (ina_kozuharova@yahoo.co.uk)

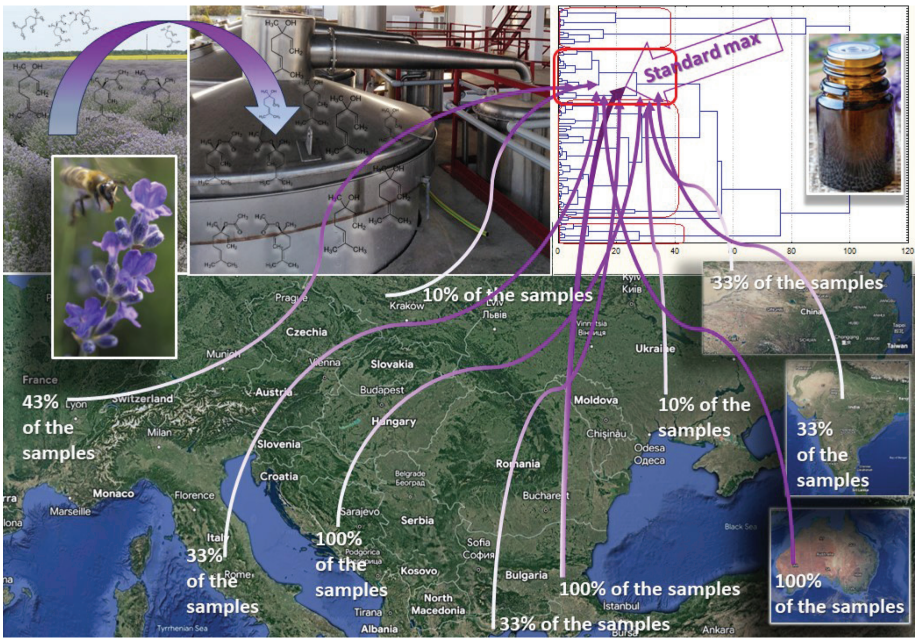
Received 11 May 2024 ♦ **Accepted** 31 May 2024 ♦ **Published** 21 June 2024

Citation: Kozuharova E, Simeonov V, Stoycheva C, Benbassat N, Batovska D (2024) Lavender essential oils–hidden relationships between the samples of origin. Pharmacia 71: 1–10. https://doi.org/10.3897/pharmacia.71.e127293

Abstract

Lavender essential oil is an economically important ingredient in perfumery, the food industry, and pharmacy. There is notable diversity in the composition of lavender essential oils. The reasons are the high genetic diversity of lavender cultivars and the variety of ecological specifics in the regions of cultivation. The aim of this research is to check which lavender essential oil variety is best regarding the content of the most important components through comparative statistical tests. We created a data set of 88 lavender essential oil samples from 16 countries. The multivariate statistics (hierarchical and non-hierarchical clustering) and factor analysis reveal hidden relationships between the objects of the study (samples) or between the variables characterizing the objects (chemical descriptors–16 components). The results are discussed in detail. All samples from Bulgaria, together with a few of the Italian, French, Greek, Indian, and Chinese samples, fall into one cluster with the standard maximums.

Graphical abstract:



PENSOFT.

Keywords

Lavandula angustifolia essential oil, multivariate statistics, hierarchical and non-hierarchical clustering (K-means clustering), factor analysis

Introduction

Lavender essential oil is among the top five in aromatherapy (Clarke 2002, 2009; Lawless 2013; Buckle 2014). Also, it is an economically important ingredient in perfumery, the food industry, and pharmacy (Lis-Balchin 2002; Ramsey et al. 2020).

Lavender essential oil can help against insomnia and anxiety. It has painkiller, anti-inflammatory, anti-allergic, and anti-microbial activity. Additionally, lavender essential oil demonstrates insect repellent and acaricidal properties. Linalool and linalyl acetate are the main ingredients of lavender essential oil. Their quantity characterizes their quality. Their quantities vary. Linalool is responsible for the sedative and painkiller effects, as well as the anti-inflammatory activity of lavender essential oil. It has antioxidant and antitumor effects, as well as antimicrobial activity. Linalyl acetate has an anti-inflammatory effect. It can prevent hypertension-related ischemic injury and the development of type 2 diabetes mellitus. Although these two components have certain pharmacological effects, the essential oil as a whole often possesses higher efficiency due to the synergism of all the constituents (Kozuharova et al. 2023).

Lavender (*Lavandula angustifolia* Mill. subsp. *angustifolia* (*L. spica* L. var. alpha, *L. officinalis* Chaix, *L. fragrans* Jord., *L. vera* DC) is one of the over 30 species of the genus *Lavandula* L. (family Lamiaceae). In terms of its general morphology, this genus is a rather mixed and divergent group of shrubs, woody-based perennials, or short-lived herbs, often aromatic, glabrous, or with a variable indumentum. *L. angustifolia* is a small woody shrub. The leaves are simple and linear-lanceolate in shape. They are gray tomentose when young, becoming greener with age. The inflorescence stalk is usually unbranched and bears a compact spike. The spike consists of cymes, and each cyme has many flowers (3–)5–7(–9). The diagnostic feature is the shape of the bracts. Those of *L. angustifolia* are broadly ovate-rhombic to obovate, unlike *L. latifolia* L., which has linear bracts. *L. angustifolia* is native to SW and South-Central Europe (Italy, France, and Spain), and it grows in the mountains (Upson, 2002). Only three members of the genus *Lavandula* are industrially cultivated for the production of essential oils: lavender (*L. angustifolia*), spike lavender (*L. latifolia*), and lavandin, a sterile hybrid developed by crossing *L. angustifolia* × *L. latifolia*. The essential oil of *L. angustifolia* is more expensive than that of other lavender species on the market because of its high quality and the plant's low yield of essential oil (Lis-Balchin 2002; Lesage-Meessen et al. 2015; Giray 2018).

Lavender is cultivated worldwide in a number of countries, and the leading lavender oil producers are Bulgaria, France, the UK, China, India, Spain, and others (Giray 2018). During the last few years, the volumes of lavender oil produced in Bulgaria steadily exceeded those of France, and the country became the world's top lavender oil producer (Giray 2018; Stanev et al. 2016; Zagorcheva et al. 2020), growing over 11145 ha of different lavender cultivars (varieties) and 50126 tons of lavender yield in 2022 (Anonymous, 2022). The high genetic diversity of the various cultivars as well as the high variety of ecological specifics in the regions of cultivation suggest a notable diversity in the composition of lavender essential oils (Giray 2018; Zagorcheva et al. 2013, 2020).

The aim of this research is to check which lavender essential oil variety is best regarding the content of the most important components through comparative statistical tests.

Material and methods

Data set preparation

We accessed Google Scholar, Web of Science, and PubMed to identify publications for the period 1900–2022, with the search string "*Lavandula angustifolia* + essential oil", "linalool", "pharmacological effects", etc. Also, we included in the search string "Bulgaria", "China", "India", and "France"– the countries that are recognized as main world producers of lavender essential oil, as well as "Greece", "Ukraine", "Poland", and other neighboring countries. Following the PRISMA 2000 guidelines, the records were assessed for eligibility, and the inappropriate ones were excluded. We selected 32 publications that presented the full components' lists as a result of GC or GC-MS analyses of lavender (*L. angustifolia*) essential oil obtained by hydrodistillation (either industrially or using Clevenger apparatus for 40 minutes to 4 hours) from various counties and varieties, under different cultivation. Based on published results of various samples (each publication presented analyses of 1 to 19 samples) was created a data set in Excel (Kozuharova et al. 2023; Naef and Morris 1992; Venskutonis et al. 1997; Milhau et al. 1997; Adam et al. 1998; Ghelardini et al. 1999; Renaud et al. 2001; Shellie et al. 2002; Chakopoulou et al. 2003; Schmidt 2003; Afsharypuor and Azarbayejany 2006; Ihsan 2006; Babu and Singh 2007; Cong et al. 2008; Smigielski et al. 2009; Hassiotis et al. 2010a, 2010b; Verma et al. 2010; Wesolowska et al. 2010; Raina and Negi 2012; Adaszyńska et al. 2013; Prusinowska and Śmigielski 2014; Konaktchiev 2015; Zagorcheva et al. 2016; Lafhal et al. 2016; Nikšić et al. 2017; Luo et al. 2019; Détár et al. 2020; Dong et al. 2020; Chen et al. 2020; Gangoo et al. 2021; Pokajewicz et al. 2021; Cong et al. 2022). Initially, it comprised 99 samples. In total, 309 names of components were used in these publications. We checked for synonyms in the compound names and eliminated the duplications (compounds were alphabetically sorted). The records of isomers and coeluates were combined and generalized since the main limitation of the published data was the lack of homogeny. Eleven samples were removed from the data set due to the lack of consistency (for example, linalool and linalyl acetate quantities were much lower than the minimum of the standard), namely samples from the USA (6), Romania (1), China (1), and unknown locations (3). Also, the components that were not detected in more than 10 of the samples were removed. The final dataset prepared for the analysis consists of 88 lavender essential oil samples and 16 components (Suppl. material 1). Additionally, we included in our data set the values of the standards of the International Organization for Standardization (ISO) and European Pharmacopoeia (10th edn) (Table 1) (Lis-Balchin 2002; ISO 3515:2002 2018; European Pharmacopoeia 2020).

Table 1. Contents of the main lavender essential oil components of the standards (ISO 3515:2002, European Pharmacopoeia 2020).

Components	ISO and/or European Pharmacopoeia (10th edition)	
	Min [%]	Max [%]
Camphor		1.20 or 1.50
1,8-Cineole		2.50 or 3.00
1,8-Cineole + Phellandrene		
Phellandrene		1.00
D-Limonene		1.00
Z-β-Ocimene	1.00	10.00
E-β-Ocimene	0.50	6.00
Lavandulol	0.1	3.00
Lavandulyl acetate	0.20	8.00
Linalool	20.00	45.00
Linalyl acetate	25.00	47.00
3-Octanone	0.10	5.00
Terpinen-4-ol	0.10	8.00
α-Terpineol		2.00

Statistical analyses

Chemometric methods were used to mine the data set of dimensions [90 × 16] (88 experimentally tested samples of lavender essential oils along with minimum and maximum values of the standards; see above) and subject it to multivariate statistical analysis. The lavender essential oil samples are from different geographical locations. Ukraine (UA, n=30), Bulgaria (BG, n=14), Poland (PL, n=10), France (FR, n=7), India (IN, n=6), Greece (GR, n=6), Italy (IT, n=3), Australia (AU, n=2), China (CN, n=3), Jordan (JO, n=2), Iran (IR, n=1), Hungary (HU, n=1), Lithuania (LT, n=1), Moldova (MD, n=1), United Kingdom (UK, n=1), and Bosnia and Herzegovina (BA, n=1) were analyzed for the content of 16 important chemical components.

The major goal of the chemometric study was to reveal hidden relationships between the objects of the study (partitioning with respect to the geographical location) and between the variables characterizing the objects (chemical descriptors). An additional task to the partitioning procedure was the determination of the specific descriptors for each group of similarity found in the partitioning procedure.

The multivariate statistical methods used for the data mining were hierarchical and non-hierarchical clustering (K-means clustering) and, additionally, factor analysis. All of the methods are well-known and described fully in the literature (Massart et al. 1987, 1998). In general, cluster analysis aims to find patterns of similarity (clusters) in a large dataset. It could happen by spontaneous clustering without preliminary conditions (unsupervised approach, hierarchical clustering) or by a supervised approach where the number of clusters is a priori determined by experts or by checking a preliminary hypothesis (K-means clustering as a supervised method). The clustering methods follow specific algorithms such as data standardization, choice of similarity measures, linkage of the similar objects into clusters, and check of the cluster significance. In hierarchical clustering, the output is a graphical plot called a dendrogram, since in K-means clustering, the results of partitioning are presented as tables describing the membership of each object or variable.

Results

Hierarchical cluster analysis

The first step of our analysis was to build the hierarchical dendrograms for linkage of the objects (Fig. 1) and variables (Fig. 2) into patterns of similarity. The exact membership of each object is presented for convenience in membership tables (Tables 2, 3) as the output of the K-means clustering (it is worth noting that the clustering results for both methods are very similar). On both dendrograms, it is quite obvious that the objects are partitioned into four groups of similarity (clusters). The chemical descriptors are separated into four clusters (Fig. 1). Also, four clusters of objects (locations of origin of the lavender essential oil samples and standards' minimum and maximum values) are formed (Fig. 2).

K-means clustering

As already mentioned above, a priory hypothesis, along with expert opinion, required partitioning of the variables and the objects into 4 clusters for each category. The members of each cluster of objects and variables are presented in Tables 2, 3. The results from K-means clustering enable the interpretation of the four clusters and the discussion of the membership of each variable.

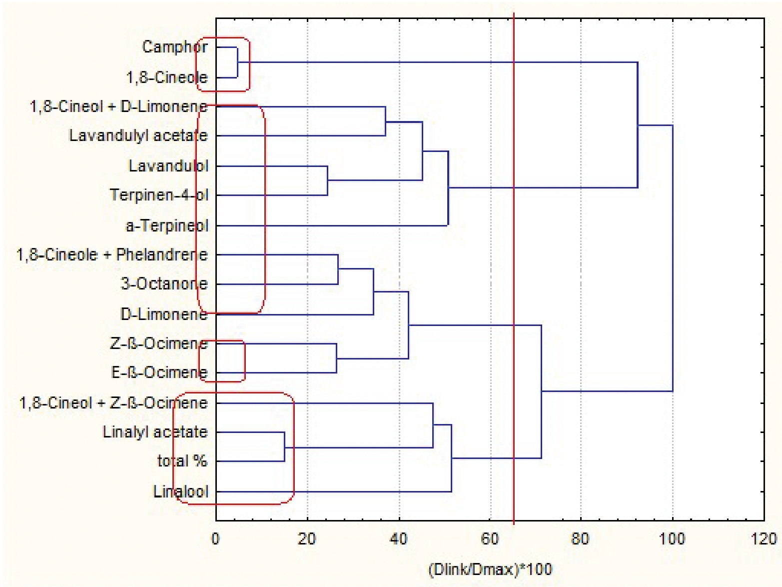


Figure 1. Hierarchical dendrogram for clustering of 16 chemical variables–the main components of the lavender essential oil.

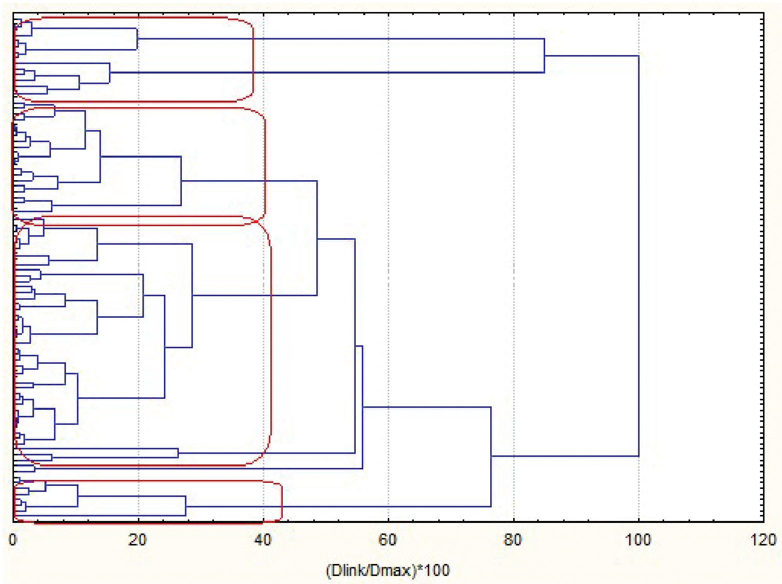


Figure 2. Hierarchical dendrogram for clustering of 90 objects (88 samples and two standards).

The four clusters of variables (Fig. 1, Table 2) give an indication of the factors responsible for the explanation of the dataset structure. It becomes possible to suggest that the important conditional components of the structure are the total explained content of the *Lavandula angustifolia* essential oil being correlated to 1,8-cineole + phellandrene, D-limonene, Z-β-ocimene, linalyl acetate, and 3-Octanone as components in lower concentrations; another two factors are related to camphor, 1,8-cineol, and ocimene derivates as aromatic components; and one factor for components in higher concentrations as a macro-component factor. Interestingly, the components linalool and linalyl acetate, responsible for the specific pharmacological effects of lavender essential oil and characterizing its quality, fall in different clusters, namely Cluster Number 1 and Cluster Number 4 (Table 2).

It is of substantial interest to find out which descriptors are specific to each one of the identified clusters. The plot of averages for each variable for each identified cluster is presented in Fig. 3. For better interpretation of the plot, below is the sequence of all variables (on the plot, 8 out of all 16 variables are presented). The results for specific descriptors of the four clusters are presented in Tables 3, 4.

The four clusters of objects (locations of origin of the lavender essential oil samples and minimum and maximum values of the standards, Fig. 1, Table 5) inform some specific geographical dependencies related to *Lavandula angustifolia* essential oil quality.

Cluster Number 1 contains 32 members (Table 5), 14 of which are lavender essential oil samples with Bulgarian origin (100% of the Bulgarian samples), and three samples of French origin (43% of the French samples),

Table 2. Membership of 16 variables into four clusters.

Variables	Distance
Members of Cluster Number 1 and Distances from Respective Cluster Center Cluster contains 6 variables	
Variables	**Distance**
1,8-Cineole + Phelandrene	0,786206
D-Limonene	0,816296
Z-β-Ocimene	0,802858
Linalyl acetate	0,746208
3-Octanone	0,744208
total %	0,832584
Members of Cluster Number 2 and Distances from Respective Cluster Center Cluster contains 2 variables	
Variables	**Distance**
Camphor	0,225288
1,8-Cineole	0,225288
Members of Cluster Number 3 and Distances from Respective Cluster Center Cluster contains 2 variables	
Variables	**Distance**
1,8-Cineol + Z-β-Ocimene	0,609882
E-β-Ocimene	0,609882
Members of Cluster Number 4 and Distances from Respective Cluster Center Cluster contains 6 variables	
Variables	**Distance**
1,8-Cineol + D-Limonene	0,818032
Lavandulol	0,742492
Lavandulyl acetate	0,906411
Linalool	0,981822
Terpinen-4-ol	0,821831
α-Terpineol	0,978135

three samples of Ukrainian origin (11% of the Ukrainian samples), two samples of Australian origin (100% of the Australian samples), two samples of Indian origin (33% of the Indian samples), two samples of Greek origin (33% of the Greek samples), one sample from China (33% of the Chinese samples), and three other single samples as follows - from Bosnia and Hercegovina, (the only one from this location), from Italy (33% of the Italian samples), and from Poland (10% of the Polish samples). It is important to note that this cluster includes the variable "standard max" (the maximum values). This is an indication that the Bulgarian samples, as well as the other listed samples, are closely bonding to the maximum values of the standard requirements. This cluster could be a conditionally named "Bulgarian" pattern.

Table 3. Specific descriptors of the four clusters. Legend: For convenience, the variables are partially numbered, e.g., camphor is number 1, D-limonene is number 5, etc.

1 Camphor	1,8-Cineole	1,8-Cineole + Phelandrene	1,8-Cineol + D-Limonene
5 D-Limonene	1,8-Cineol + Z-β-Ocimene	Z-β-Ocimene	E-β-Ocimene
9 Lavandulol	Lavandulyl acetate	Linalool	Linalyl acetate
13 3-Octanone	Terpinen-4-ol	α-Terpineol	Total %

Table 4. Specific descriptors (highest and lowest levels) for all identified clusters.

Clusters	Highest levels of	Lowest levels of
1	1,8-Cineol+z-β-Ocimene 1,8-Cineol+Phelandrene E-β-Ocimene Linalyl acetatete 3- Octanone Total	Terpineol
2	–	E-β-Ocimene 1,8-Cineol+Z-β-Ocimene
3	Lavandulol Terpinen-4-ol, α-Terpineol	Linalyl acetatete Total
4	Camphor 1,8-Cineole	Linalool

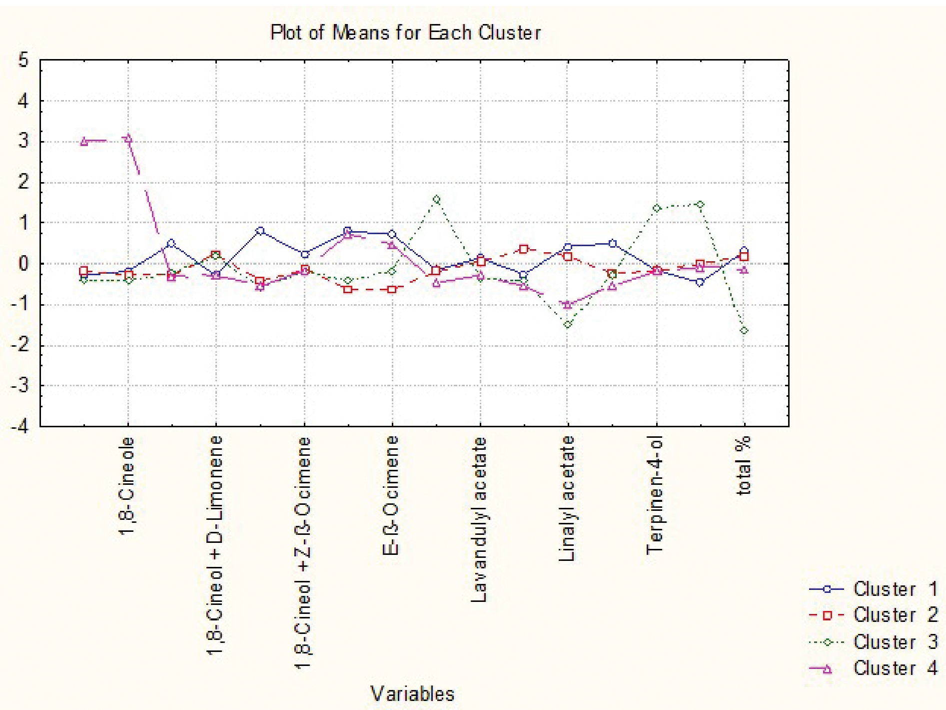


Figure 3. Plot of means for each descriptor for each identified cluster.

Table 5. Membership of 90 objects into four clusters.

Members of Cluster Number 1 and Distances from Respective Cluster Center Cluster contains 32 cases	
Country of origin of the essential oils	**Distance**
PL-2	0,803600
BG-5-1	1,063737
FR-5-3	1,131704
BA-8	0,841112
AU-9	1,099822
MD-10-2	0,492620
UA-10-3	0,498780
BG-10-4	0,512807
AU-10-5	0,876861
IT-14	0,771338
BG-16-1	1,073207
BG-16-2	0,590999
BG-16-3	1,420205
BG-16-4	0,786706
BG-16-5	0,865210
BG-16-6	0,671406
BG-16-7	0,899521
BG-17-1	0,572984
BG-17-2	0,546892
BG-17-3	0,535056
GR-18-1	1,034623
UK-21-6	0,882925
GR-24	0,778940
UA-28-3	0,779688
BG-29	0,537594
CN-31	0,888515
FR-33-1	1,337309
FR-33-3	1,984810
IN-35-1	0,728521
IN-35-3	0,702783
BG-36-1	0,733750
standart max	1,656808
Members of Cluster Number 2 and Distances from Respective Cluster Center Cluster contains 41 cases	
Country of origin of the essential oils	**Distance**
Pl-1-2	0,740580
CN-4	0,985789
IT-5-2	0,709475
PL-5-4	0,741186
IT-6	0,694345
IN-7	0,594857
FR-10-1	0,449550
LT-12	0,707037
FR-13	0,661259
HU-21-3	0,962680
UK-21-4	0,691624
FR-21-5	0,674496
IN-26	0,547958
UA-27-1	0,533303
UA-27-2	0,513843
UA-27-3	0,782405
UA-27-4	0,714650
UA-27-5	2,019227
UA-27-6	0,368959
UA-27-7	0,656498
UA-27-8	0,708169
UA-28-1	0,669752
UA-28-2	0,421578
UA-28-4	0,564641
Members of Cluster Number 2 and Distances from Respective Cluster Center Cluster contains 41 cases	
Country of origin of the essential oils	**Distance**
UA-28-5	0,474619
UA-28-6	0,717241
UA-28-7	0,278719
UA-28-8	0,257587
UA-28-9	0,270680
UA-28-10	0,227829
UA-28-11	0,579047
UA-28-12	0,428956
UA-28-13	0,570951
UA-28-15	0,511818
UA-28-16	0,387688
UA-28-17	0,956365
CN-30	0,480355
FR-33-2	0,738944
IN-34	0,744650
IN-35-2	1,285159
standart min	0,641462
Members of Cluster Number 3 and Distances from Respective Cluster Center Cluster contains 10 cases	
Country of origin of the essential oils	**Distance**
PL-1-1	0,979280
PL-1-3	0,800925
PL-22-1	0,601253
PL-22-2	0,532358
PL-22-3	0,803043
PL-22-4	0,681577
PL-22-5	0,591797
UA-28-14	1,212039
UA-28-18	1,398853
UA-28-19	0,899213
Members of Cluster Number 4 and Distances from Respective Cluster Center Cluster contains 7 cases	
Country of origin of the essential oils	**Distance**
GR-15	1,166784
GR-18-2	0,494226
GR-19-1	0,584855
GR-19-2	0,750952
IR-23	0,876077
JO-25-1	0,521416
JO-25-2	0,511501

Cluster Number 2 includes 41 members and is the biggest one (Table 5). It includes predominantly samples from Ukraine (23 out of a total of 41 samples, or 56%). This is also 82% of all Ukrainian samples. Next are four samples of French origin (57% of the French samples), followed by four samples of Indian origin (67% of the Indian samples), two samples from Poland (20% of the Polish samples), two samples from Italy (67% of the Italian samples), two samples from China (67% of the Chinese samples), two samples from the United Kingdom (100% of the UK samples), and two single samples (each of them, the only one from this location) as follows from Hungary and Lithuania. Again, a standard sample (the minimum values) belongs to this cluster, and it allows the assumption to name this cluster the "Ukrainian" pattern conditionally.

Analogically, Cluster Number 3 consists of 10 members (Table 5) and represents a "Polish" pattern since the samples from Poland are seven (out of a total of 10 samples from Poland, or 70% of all samples in this cluster and 70% of all Polish samples). It is not a surprise that the remaining three samples are from Ukraine (10% of the Ukrainian samples), which is a geographical neighbor of Poland.

Finally, Cluster Number 4 is the smallest one with only seven members (Table 5), with dominantly Greek samples (4 out of total 7, and 66% of all Greek samples), and the cluster could be a conditionally named "Greek" pattern.

Factor analysis

Factor analysis (often named principal components analysis, or PCA) is a typical projection method. It helps to clarify the dataset structure by reducing variables and replacing the initial variables with latent variables (factors). The new factors are characterized by factor loadings (indicating relationships between the initial variables) and factor scores (representing the new special coordinates of the objects). The output is usually a specific table with factor loadings, which helps to interpret the physical meaning of the latent factors. In Table 6, the factor loadings for four factors are presented.

Table 6. Factor loadings.

Factor Loadings (Varimax normalized) Extraction: Principal components (marked loadings are >.700000)				
Variables	**Factor - 1**	**Factor - 2**	**Factor - 3**	**Factor - 4**
Camphor	0,114	**-0,935**	0,057	-0,002
1,8-Cineole	0,124	**-0,951**	0,094	-0,019
1,8-Cineole + Phelandrene	0,079	0,271	**0,709**	0,133
1,8-Cineol + D-Limonene	-0,448	0,059	-0,378	0,186
D-Limonene	0,055	0,020	**0,768**	0,147
1,8-Cineol + Z-β-Ocimene	0,032	0,103	0,076	0,125
Z-β-Ocimene	0,055	-0,129	**0,778**	0,011
E-β-Ocimene	-0,013	-0,139	**0,796**	0,024
Lavandulol	**-0,858**	0,093	-0,029	-0,243
Lavandulyl acetate	**-0,715**	0,083	0,108	0,223
Linalool	-0,105	0,017	**-0,702**	0,410
Linalyl acetate	0,450	0,408	0,241	**0,774**
3-Octanone	-0,285	0,192	**0,781**	0,286
Terpinen-4-ol	**-0,727**	-0,009	-0,050	-0,150
α-Terpineol	0,127	0,088	-0,131	**-0,786**
total %	0,195	0,037	-0,011	**0,881**
Expl. Var. %	*21,8*	*17,7*	*16,6*	*14,1*

The first latent factor explains nearly 22% of the total variance of the system. It indicates the role of lavandulol, lavandulyl acetate, and terpinen-4-ol on the data structure and could be a conditionally named "terpene effect" factor, which characterizes Polish samples (Table 6).

The second latent factor explains another 18% of the total variance and is related to the high loadings of camphor and 1,8-cineole being very important characteristics of the lavender oil. The Greek and some Mediterranean samples are related to this conditional "Camphor" factor (Table 6).

The next latent factor, with an explanation of almost 17% of the total variance, is related to the special derivates of ocimene and limonene. The third latent factor could conditionally be named an "aromatic" factor (Table 6).

The last latent factor 4 (about 14% of the total variance) indicates the significance of α-terpineol and linalyl acetate as major components to form the total content of the sample quantity; conditionally, this factor could be known as the "macro components" factor (Table 6).

Discussion

The ISO defines lavender essential oil as the "oil obtained by steam distillation of recently cut flowering tops of *Lavandula angustifolia* Mill." and fixes the main chemical components, which should be within a certain range of content [percents] (ISO 3515:2002 2018). The values recommended by the European Pharmacopoeia, 10th edn (European Pharmacopoeia 2020) are slightly different concerning some of the components (Table 1). The conventional method for the quality assessment of lavender oil is gas chromatography, sometimes combined with mass spectrometry with a flame ionization detector. Improvements for quality evaluations have been offered. For example, vibrational spectroscopy methods such as mid-infrared and near-infrared are combined with chemometric data analysis (Tankeu et al. 2014). Later on, it is shown that the quality of the lavender essential oil is assessed not only by the high percentage abundance of linalool and linalyl acetate, respectively, and a low percentage abundance of camphor and 1,8-cineole, but also by the concentration of the entire product. A chemometric model is offered for quality evaluation, motivating that a simple GC/MS analysis to calculate the percentage abundance of the compounds is not enough (Marincaş and Feher 2018). Another approach proposes the Q-Index method of analysis for the evaluation of lavender essential oil quality. A comparative analysis is performed based on a large set of lavender essential oil samples (n = 72) using multiple techniques (GC/MS, GC/Q-ToF, NMR, chemometric analysis, and the Q-Index method) (Wang et al. 2021). Our study suggests an alternative approach using multivariate statistics (hierarchical and non-hierarchical clustering) and factor analysis, which reveal hidden relationships between the objects of the study (samples) or between the variables characterizing the objects (chemical descriptors–16 components). The main highlights of our findings are that cluster number 1 indicates high quality because the values of standard maximums fall in it. It contains all (100%, n=14) of the Bulgarian samples, together with a few of the French, Greek, Indian, Chinese, etc. samples (see above), although the content of neither linalool nor linalyl acetate is highest in the Bulgarian samples (Suppl. material 1). Linalool and linalyl acetate largely contribute to the anxiolytic, sedative, antioxidant, anti-inflammatory, antibacterial, etc. effects attributed to lavender essential oil, but the overall efficacy is due to the synergism of the components (Kozuharova et al. 2023). Linalyl acetate and α-terpineol are pointed out by the factor analysis of our sample set (n=88, Table 6) as major components to form the total content of the sample quantity. Also, according to the ISO (ISO 3515:2002 2018) standard, lavender essential oil should contain camphor (0.5–1.0%), while the camphor content could reach 6–8% in lavandin essential oil (obtained from *Lavandula intermedia* Emeric ex Loisel, syn. *L. hybrid*, which is a hybrid of *L. angustifolia* and *L. latifolia*) (Kıvrak, 2018). The Greek and some of the other Mediterranean samples are related to the "camphor" latent factor. The content of camphor in these samples is rather high, with values close to those appointed for lavandin essential oil (Suppl. material 1).

Conclusion

Based on a large data set, we used multivariate statistical methods. The data mining was hierarchical and non-hierarchical clustering (K-means clustering) and, additionally, factor analysis in order to reveal hidden relationships between the objects of the study (partitioning with respect to the geographical location) or between the variables characterizing the objects (chemical descriptors). On both hierarchical dendrograms for linkage of the objects and variables into patterns of similarity, it is quite obvious that the objects are partitioned into 4 groups of similarity (clusters). K-means clustering is used to discuss the priory hypothesis along with expert opinion. Attention deserves Cluster Number 1, which indicates high quality. It contains 32 members (31 samples from various countries and standard maximums). This cluster could be conditionally named the "Bulgarian" pattern because it contains all the samples from Bulgaria together with a few of the French, Greek, Indian, Chinese, etc. samples. The applied factor analysis explains the total variance of the system and confirms the significance of the latent components (conditionally named "terpene" and "camphor" as very specific descriptors for the overall data structure further; "aromatic" and "macrocomponent" factors as general indicators for the lavender essential oil quality). Thus, the results from the statistical data mining could be of use for predicting the quality of lavender essential oil if a limited number of chemical components are available.

Acknowledgements

Ekaterina Kozuharova was supported in this research by the European Union, NextGenerationEU, through the National Recovery and Resilience Plan of the Republic of Bulgaria, Project № BG-RRP-2.004-0004-C01.

References

Adam K, Sivropoulou A, Kokkini S, Lanaras T, Arsenakis M (1998) Antifungal activities of *Origanum vulgare* subsp. *hirtum*, *Mentha spicata*, *Lavandula angustifolia*, and *Salvia fruticosa* essential oils against human pathogenic fungi. Journal of Agricultural and Food Chemistry 46(5): 1739–1745. https://doi.org/10.1021/jf9708296

Adaszyńska M, Swarcewicz M, Dzięcioł M, Dobrowolska A (2013) Comparison of chemical composition and antibacterial activity of lavender varieties from Poland. Natural Product Research 27(16): 1497–1501. https://doi.org/10.1080/14786419.2012.724408

Afsharypuor S, Azarbayejany N (2006) Chemical constituents of the flower essential oil of *Lavandula officinalis* Chaix. from Isfahan (Iran). Iranian Journal of Pharmaceutical Sciences 2(3): 169–172.

Anonymous (2022) Operational analysis for major agricultural crops 30. Ministry of Agriculture of Bulgaria (in Bulgarian). https://www.mzh.government.bg/media/filer_public/2022/07/28/operativen_analiz_2022-07-27_ZiA805x.pdf [Accessed on 19.08.2023]

Babu GK, Singh B (2007) Comparative chemical composition of direct, recovered and combined essential oils of *Lavandula angustifolia* Mill. India Perfumers 51(4): 50.

Buckle J (2014) Clinical Aromatherapy-E-Book: Essential Oils in Healthcare. Elsevier Health Sciences, 412 pp. https://doi.org/10.1016/B978-0-7020-5440-2.00002-4

Chakopoulou PS, Goliaris AH, Katsiotis ST (2003) Contribution to the analysis of the volatile constituents from some lavender and lavandin cultivars grown in Greece. Scientia Pharmaceutica 71(3): 229–234. https://doi.org/10.3797/scipharm.aut-03-21

Chen X, Zhang L, Qian C, Du Z, Xu P, Xiang Z (2020) Chemical compositions of essential oil extracted from *Lavandula angustifolia* and its prevention of TPA-induced inflammation. Microchemical Journal 153: 104458. https://doi.org/10.1016/j.microc.2019.104458

Clarke S (2002) Essential chemistry for safe aromatherapy. Churchill Livingstone, Edinburgh, 231 pp.

Clarke S (2009) Essential Chemistry for Aromatherapy E-Book. Elsevier Health Sciences, 301 pp.

Cong K, Białoń M, Svydenko L, Hudz N, Balwierz R, Marciniak D, Wieczorek PP (2022) Comparative evaluation of the essential oil of the new Ukrainian *Lavandula angustifolia* and *Lavandula × intermedia* cultivars grown on the same plots. Molecules 27(7): 2152. https://doi.org/10.3390/molecules27072152

Cong Y, Abulizi P, Zhi L, Wang X, Mirensha M (2008) Chemical composition of the essential oil of *Lavandula angustifolia* from Xinjiang, China. Chemistry of Natural Compounds 44: 810–810. https://doi.org/10.1007/s10600-009-9210-8

Détár E, Németh ÉZ, Gosztola B, Demján I, Pluhár Z (2020) Effects of variety and growth year on the essential oil properties of lavender (*Lavandula angustifolia* Mill.) and lavandin (*Lavandula × intermedia* Emeric ex Loisel.). Biochemical Systematics and Ecology 90: 104020. https://doi.org/10.1016/j.bse.2020.104020

Dong G, Bai X, Aimila A, Aisa HA, Maiwulanjiang M (2020) Study on lavender essential oil chemical compositions by GC-MS and improved pGC. Molecules 25(14): 3166. https://doi.org/10.3390/molecules25143166

European Pharmacopoeia (2020) 10th edn.; Council of Europe: Strasbourg, France.

Gangoo SA, Malik AR, Peerzada I, Amarjeet S (2021) Research Article Advances in Agriculture, Horticulture and Entomology AAHE-161 ISSN 2690-1900.

Ghelardini C, Galeotti N, Salvatore G, Mazzanti G (1999) Local anaesthetic activity of the essential oil of *Lavandula angustifolia*. Planta Medica 65(08): 700–703. https://doi.org/10.1055/s-1999-14045

Giray FH (2018) An analysis of world lavender oil markets and lessons for Turkey. Journal of Essential Oil Bearing Plants 21(6): 1612–1623. https://doi.org/10.1080/0972060X.2019.1574612

Hassiotis CN, Lazari DM, Vlachonasios KE (2010a) The effects of habitat type and diurnal harvest on essential oil yield and composition of *Lavandula angustifolia* Mill. Fresenius Environmental Bulletin 19(8): 1491–1498.

Hassiotis CN, Tarantilis PA, Daferera D, Polissiou MG (2010b) Etherio, a new variety of *Lavandula angustifolia* with improved essential oil production and composition from natural selected genotypes growing in Greece. Industrial Crops and Products 32(2): 77–82. https://doi.org/10.1016/j.indcrop.2010.03.004

Ihsan AS (2006) Essential oil composition of *Lavandula officinalis* L. grown in Jordan. Journal of University of Kerbala 2(3): 81–21.

ISO 3515:2002 (2018) Oil of lavender (*Lavandula angustifolia* Mill.) This standard was last reviewed and confirmed in 2018. https://www.iso.org/standard/36253.html [Accessed on 11.08.2023]

Kıvrak Ş (2018) Essential oil composition and antioxidant activities of eight cultivars of Lavender and Lavandin from western Anatolia. Industrial Crops and Products 117: 88–96. https://doi.org/10.1016/j.indcrop.2018.02.089

Konaktchiev A (2015) Essential oils from *Lavandula angustifolia* Mill. varieties and species of the genus *Achillea* L. PhD Thesis, IOCCPC, BAS, Sofia. [In Bulgarian]

Kozuharova E, Simeonov V, Batovska D, Stoycheva C, Valchev H, Benbassat N (2023) Chemical composition and comparative analysis of lavender essential oil samples from Bulgaria in relation to the pharmacological effects. Pharmacia 70(2): 395–403. https://doi.org/10.3897/pharmacia.70.e104404

Lafhal S, Vanloot P, Bombarda I, Kister J, Dupuy N (2016) Chemometric analysis of French lavender and lavandin essential oils by near infrared spectroscopy. Industrial Crops and Products 80: 156–164. https://doi.org/10.1016/j.indcrop.2015.11.017

Lawless J (2013) The encyclopedia of essential oils. Conari Press, San Francisco, 224 pp.

Lesage-Meessen L, Bou M, Sigoillot JC, Faulds CB, Lomascolo A (2015) Essential oils and distilled straws of lavender and lavandin: a review of current use and potential application in white biotechnology. Applied Microbiology and Biotechnology 99(8): 3375–3385. https://doi.org/10.1007/s00253-015-6511-7

Lis-Balchin M (2002) Lavender essential oil Standardisation, ISO; adulteration and its detection using GC, enantiomeric columns and bioactivity. In: Lis-Balchin M (Ed.) Lavender: The Genus *Lavandula*. CRC Press, London, 131–137. https://doi.org/10.1201/9780203216521-16

Luo W, Du Z, Zheng Y, Liang X, Huang G, Zhang Q, Liu Z, Zhang K, Zheng X, Lin L, Zhang L (2019) Phytochemical composition and bioactivities of essential oils from six Lamiaceae species. Industrial Crops and Products 133: 357–364. https://doi.org/10.1016/j.indcrop.2019.03.025

Marincaş O, Feher I (2018) A new cost-effective approach for lavender essential oils quality assessment. Industrial Crops and Products 125: 241–247. https://doi.org/10.1016/j.indcrop.2018.09.010

Massart DL, Kaufman L (1987) The Interpretation of Analytical Chemical Data by the Use of Cluster Analysis. J. Wiley & Sons Incorp., John Wiley & Sons, New York.

Massart DL, Vandeginste BG, Buydens LM, Lewi PJ, Smeyers-Verbeke J, Jong SD (1998) Handbook of chemometrics and qualimetrics. Elsevier Science Inc., Amsterdam.

Milhau G, Valentin A, Benoit F, Mallié M, Bastide JM, Pélissier Y, Bessière JM (1997) In vitro antimalarial activity of eight essential oils. Journal of Essential Oil Research 9(3): 329–333. https://doi.org/10.1080/10412905.1997.10554252

Naef R, Morris AF (1992) Lavender–Lavandin. A comparison. Riv. Ital. EPPOS, 3, 364–377.

Nikšić H, Kovač-Bešović E, Makarević E, Durić K, Kusturica J, Muratovic S (2017) Antiproliferative, antimicrobial, and antioxidant activity of *Lavandula angustifolia* Mill. essential oil. Journal of Health Sciences 7: 35–43. https://doi.org/10.17532/jhsci.2017.412

Pokajewicz K, Białoń M, Svydenko L, Fedin R, Hudz N (2021) Chemical composition of the essential oil of the new cultivars of *Lavandula angustifolia* Mill. Bred in Ukraine. Molecules 26(18): 5681. https://doi.org/10.3390/molecules26185681

Prusinowska R, Śmigielski KB (2014) Composition, biological properties and therapeutic effects of lavender L). A review. Herba Polonica 60(2): 56–66. https://doi.org/10.2478/hepo-2014-0010

Raina AP, Negi KS (2012) Comparative essential oil composition of *Lavandula* species from India. Journal of Herbs, Spices & Medicinal Plants 18(3): 268–273. https://doi.org/10.1080/10496475.2012.690142

Ramsey JT, Shropshire BC, Nagy TR, Chambers KD, Li Y, Korach KS (2020) Essential oils and health. Yale Journal of Biology and Medicine 93(2):291–305.

Renaud EN, Charles DJ, Simon JE (2001) Essential oil quantity and composition from 10 cultivars of organically grown lavender and lavandin. Journal of Essential Oil Research 13(4): 269–273. https://doi.org/10.1080/10412905.2001.9699691

Schmidt E (2003) The characteristics of Lavender oils from Eastern Europe. Perfumer and Flavorist 28(4): 48–61.

Shellie R, Mondello L, Marriott P, Dugo G (2002) Characterisation of lavender essential oils by using gas chromatography–mass spectrometry with correlation of linear retention indices and comparison with comprehensive two-dimensional gas chromatography. Journal of Chromatography A 970(1–2): 225–234. https://doi.org/10.1016/S0021-9673(02)00653-2

Smigielski K, Raj A, Krosowiak K, Gruska R (2009) Chemical composition of the essential oil of *Lavandula angustifolia* cultivated in Poland. Journal of Essential Oil Bearing Plants 12(3): 338–347. https://doi.org/10.1080/0972060X.2009.10643729

Stanev S, Zagorcheva T, Atanassov I (2016) Lavender cultivation in Bulgaria-21st century developments, breeding challenges and opportunities. Bulgarian Journal of Agricultural Science 22(4): 584–590.

Tankeu SY, Vermaak I, Kamatou GP, Viljoen AM (2014) Vibrational spectroscopy and chemometric modeling: An economical and robust quality control method for lavender oil. Industrial Crops and Products 59: 234–240. https://doi.org/10.1016/j.indcrop.2014.05.005

Upson T (2002) The taxonomy of the genus *Lavandula* L. Lavender: the genus *Lavandula*. Taylor & Francis, London, 2–34.

Venskutonis PR, Dapkevicius A, Baranauskiene M (1997) Composition of the essential oil of Lavender (*Lavandula angustifolia* Mill.) from Lithuania. Journal of Essential Oil Research 9(1): 107–110. https://doi.org/10.1080/10412905.1997.9700727

Verma RS, Rahman LU, Chanotiya CS, Verma RK, Chauhan A, Yadav A, Anand S, Yadav AK (2010) Essential oil composition of *Lavandula angustifolia* Mill. cultivated in the mid hills of Uttarakhand, India. Journal of the Serbian Chemical Society 75(3): 343–348. https://doi.org/10.2298/JSC090616015V

Wang M, Zhao J, Ali Z, Avonto C, Khan IA (2021) A novel approach for lavender essential oil authentication and quality assessment. Journal of Pharmaceutical and Biomedical Analysis 199: 114050. https://doi.org/10.1016/j.jpba.2021.114050

Wesolowska A, Jadczak D, Grzeszczuk M (2010) Influence of distillation time on the content and composition of essential oil isolated from lavender (*Lavandula angustifolia* Mill.). Herba Polonica 56(3).

Zagorcheva T, Stanev S, Rusanov K, Atanassov I (2013) Comparative GC/MS analysis of lavender (*Lavandula angustifolia* Mill.) inflorescence and essential oil volatiles. Agricultural Science and Technology 5(4): 1313–8820.

Zagorcheva T, Rusanov K, Stanev S, Atanassov I (2016) A simple procedure for comparative GC-MS analysis of lavender (*Lavandula angustifolia* Mill.) flower volatile composition. IOSR Journal of Pharmacy and Biological Sciences 11(4): 9–14. https://doi.org/10.9790/3008-1104030914

Zagorcheva T, Stanev S, Rusanov K, Atanassov I (2020) SRAP markers for genetic diversity assessment of lavender (*Lavandula angustifolia* Mill.) varieties and breeding lines. Biotechnology & Biotechnological Equipment 34(1): 303–308. https://doi.org/10.1080/13102818.2020.1742788

Supplementary material 1

Components identified in the samples from 88 lavender essential oil samples along with minimum and maximum values of the standards

Authors: Ekaterina Kozuharova, Vasil Simeonov, Christina Stoycheva, Niko Benbassat, Daniela Batovska

Data type: docx

Explanation note: Legend: Abbreviation of the country, number of the publication [1–35] and number of samples referred in these publications.

Link: https://doi.org/10.3897/pharmacia.71.e127293.suppl1