OPEN ACCESS

Comprehensive overview of common e-liquid ingredients and how they can be used to predict an e-liquid's flavour category

Erna J Z Krüsemann ,[1,2] Anne Havermans ,[1] Jeroen L A Pennings,[1] Kees de Graaf,[2] Sanne Boesveldt,[2] Reinskje Talhout[1]

► Additional material is published online only. To view please visit the journal online (http://dx.doi.org/10.1136/tobaccocontrol-2019-055447).

[1]Centre for Health Protection, Rijksinstituut voor Volksgezondheid en Milieu, Bilthoven, The Netherlands
[2]Division of Human Nutrition and Health, Wageningen University, Wageningen, The Netherlands

Correspondence to
Erna J Z Krüsemann, Centre for Health Protection, Rijksinstituut voor Volksgezondheid en Milieu, Bilthoven 3721 MA, Netherlands;
erna.krusemann@rivm.nl

EJZK and AH contributed equally.

Received 10 October 2019
Revised 10 January 2020
Accepted 19 January 2020
Published Online First 10 February 2020

To cite: Krüsemann EJZ, Havermans A, Pennings JLA, *et al*. *Tob Control* 2021;**30**:185–191.

ABSTRACT

Objectives Flavours increase e-cigarette attractiveness and use and thereby exposure to potentially toxic ingredients. An overview of e-liquid ingredients is needed to select target ingredients for chemical analytical and toxicological research and for regulatory approaches aimed at reducing e-cigarette attractiveness. Using information from e-cigarette manufacturers, we aim to identify the flavouring ingredients most frequently added to e-liquids on the Dutch market. Additionally, we used flavouring compositions to automatically classify e-liquids into flavour categories, thereby generating an overview that can facilitate market surveillance.

Methods We used a dataset containing 16 839 e-liquids that were manually classified into 16 flavour categories in our previous study. For the overall set and each flavour category, we identified flavourings present in more than 10% of the products and their median quantities. Next, quantitative and qualitative ingredient information was used to predict e-liquid flavour categories using a random forest algorithm.

Results We identified 219 unique ingredients that were added to more than 100 e-liquids, of which 213 were flavourings. The mean number of flavourings per e-liquid was 10±15. The most frequently used flavourings were vanillin (present in 35% of all liquids), ethyl maltol (32%) and ethyl butyrate (28%). In addition, we identified 29 category-specific flavourings. Moreover, e-liquids' flavour categories were predicted with an overall accuracy of 70%.

Conclusions Information from manufacturers can be used to identify frequently used and category-specific flavourings. Qualitative and quantitative ingredient information can be used to successfully predict an e-liquid's flavour category, serving as an example for regulators that have similar datasets available.

INTRODUCTION

Electronic cigarette (e-cigarette) use among various user groups has increased considerably over the past years.[1 2] One of the most important reasons for e-cigarettes' great popularity is the assortment of available e-liquid flavours[3–6]; for example, no less than 245 unique flavours were available in the Netherlands in 2017.[7] Flavours increase product attractiveness among all types of (potential) users, that is, among youth and adults[8] and among current smokers, dual users, exclusive vapers, as well as never-users.[9] For smokers, switching to e-cigarettes may be beneficial, as e-cigarette use (ie, vaping) is considered less harmful than regular cigarette smoking.[10–14] In line with this, the use and marketing of e-liquid flavours that are appealing to smokers may contribute to public health benefits. However, flavours may also stimulate vaping among non-users, in particular young people.[15–17] This is concerning, as e-cigarettes are not safe[10 18 19] That is, chemicals in e-cigarette emissions (eg, tobacco-specific nitrosamines, metals, aldehydes and other flavourings) can be toxic and thus harmful to consumers' health.[20–22] In addition, e-cigarettes may facilitate smoking initiation among never-smokers.[23] As a consequence, e-liquid flavours are considered an important target in tobacco control in order to decrease e-cigarette attractiveness and use and thereby exposure to potentially toxic emissions.

One way of regulating e-liquid flavours could be restriction of flavour categories that are particularly appealing to non-users or youth. For example, the US Food and Drug Administration (FDA) recently announced that e-cigarettes with a flavour other than tobacco will be removed from the market until they are evaluated under the FDA's new product review authority.[24] Such rules often rely on flavour-related marketing descriptions, which do not necessarily reflect human perception and may differ between e-liquid providers and therefore can be challenging to enforce. In addition, allowing or banning a specific e-liquid flavour may be complicated as e-liquids regularly have multiple flavour descriptors (eg, tobacco with caramel and vanilla).[7] Similarly, using sensory analysis to enforce e-liquid flavour regulations may be challenging. Although based on human perception, analysing sensory properties of all e-liquids in a particular country to determine whether they comply with current regulations is extremely time-consuming. However, sensory analysis could be used to determine attractiveness of particular e-liquid flavours or flavour categories among vulnerable user groups (eg, non-smoking adolescents) and in that way inform policy makers on how to reduce e-liquid appeal. Another option would be to decrease e-liquid attractiveness or toxicity by restricting the addition of particular flavouring ingredients. For example, particular flavourings could be banned or their maximum concentration could be restricted. This may diminish e-cigarette use and thereby exposure to potentially toxic e-liquid constituents and thus increase health benefits for non-users and youth.

Research on e-liquid flavours and flavouring compositions can support regulators in developing

policy measures. Accordingly, various chemical analytical studies evaluated e-liquid flavouring ingredients and emissions.[21 25–29] These studies typically focus on a list of a priori selected target flavourings for their analyses.[21 25–29] These target lists are usually selected based on previous studies, which creates a risk of selection bias by overlooking other or new e-liquid ingredients that may have toxic properties. Therefore, there is a clear need for a published overview of common flavourings and other ingredients in e-liquids. This information may be used as a starting point for future chemical analytical researchers in developing their target lists, may provide targets for future toxicological studies and may provide foci for regulation of e-liquid flavourings.

We previously generated an overview of all e-liquid flavour descriptions reported to be marketed in the Netherlands in 2017,[7] by manually classifying almost 20 000 e-liquids into 16 main flavour categories.[30] Classification was based on e-liquids' flavour-related information reported by manufacturers. In the European Union, manufacturers are also required to provide information about their e-liquids' chemical composition. Using this information, the current study first aims to identify the most commonly used e-liquid flavourings in general and to determine potential flavourings that are specific to a single flavour category.

In our previous study, we used information from manufacturers such as brand names to manually classify e-liquids by flavour description.[7] However, 2586 e-liquids (15% of the entire dataset) could not be classified as flavour-related information was unspecific, incomplete or even unavailable. For example, it was not possible to classify e-liquids with generic brand names that are unrelated to a flavour (eg, 'Spaceship' or 'Purple Unicorn', hypothetically). Manufacturers additionally reported information about their e-liquids' chemical composition. Therefore, the second aim of this study is to determine whether this information can be used to predict e-liquids' flavour categories, using a machine learning algorithm. Such automatic classification of e-liquids by flavour would allow to easily create market overviews of e-liquid flavour descriptions worldwide in a time-effective and cost-effective manner, provided that information on e-liquid ingredients is available.

METHODS

Data collection and preparation

According to the European Tobacco Product Directive, tobacco and e-liquid manufacturers are required to provide information such as brand names, ingredients and emissions of the products they have marketed in each Member State. A complete dataset of all e-cigarette products on the Dutch market was extracted from the European Common Entry Gate system (EU-CEG)[31] on 30 November 2017. For this study, only e-liquids were included (ie, no other products, such as devices). Duplicate submissions and products with incomplete information were excluded, resulting in a dataset of 19 266 products.

In a previous study, flavour-related information about each e-liquid was obtained from the EU-CEG system. According to a standardised approach, e-liquid flavours were classified into one of the following 16 main flavour categories: *tobacco*, *menthol/mint*, *nuts*, *spices*, *coffee/tea*, *alcohol*, *other beverages*, *fruit-berries*, *fruit-citrus*, *fruit-tropical*, *fruit-other*, *dessert*, *candy*, *other sweets*, *other flavours* and *unflavoured*.[7 30] E-liquids were considered 'unclassifiable' if they were not found on the internet (n=1680; 9% of total sample) or could not be searched for due to the EU-CEG information being too general (eg, only referring to a brand or product range, n=906; 5% of total sample).[7]

Data analyses

As most e-liquids were reported as being marketed in a package unit of '1' and containing 10 mL of fluid, deviating submissions (more than one e-liquid per package or e-liquids with a volume other than 10 mL) were excluded. This resulted in a final dataset of 16 839 products. For these products, ingredient-related information was extracted from the EU-CEG system and analysed using R statistical software (version 3.5.1). Ingredients reported by manufacturers as having the function 'Flavour and/or Taste Enhancer' will be referred to as flavourings. Negative values for ingredient amounts (resulting from EU-CEG artefacts) were set at zero. For the overall dataset, as well as for individual flavour categories and the unclassifiable subset separately, the following values were determined: the number of products, the mean number of total ingredients, the mean number of flavourings per product, the mean total number of ingredients per product and the mean quantities of all flavourings per product.

Ingredients present under multiple names in the EU-CEG system (eg, ethanol, etanolo, etanol, ethyl alcohol, ethyl alkohol, ethyl-alcohol, alcool ethylique and EtOH) were merged into one ingredient name. First, unique ingredient names were identified (n=8352), including the number of products for which they were reported in the EU-CEG system. Next, starting with the most frequently reported ingredients names, we manually searched for other names that represented the same ingredient and thus could be grouped together. This was done using CAS registry numbers (ie, assigned by the Chemical Abstracts Service), FEMA registry numbers (ie, assigned by the Flavor Extract Manufacturers Association), trivial names, translations of ingredient names and text cleanup (eg, removing upper/lower case redundancy and spelling mistakes). This process was repeated until all ingredient names that were initially reported in more than 100 products (ie, more than 0.6% of all products) were checked. This resulted in a final list of 219 unique ingredients.

For further analysis, solvents (propylene glycol, glycerol, ethanol, water and triacetin) and nicotine were excluded. For the remaining 213 flavourings, we identified the flavourings that were present in at least 10% of all products (n=25 flavourings), as well as the median quantity (mg/10 mL) in which they were added. This was also done for each individual flavour category and for the set of unclassifiable products (n=94 flavourings in total).

Next, quantitative information of the flavourings that were present in at least 10% of the products in any flavour category were used for machine learning prediction of an e-liquid's class (ie, flavour category) using the random forest (RF) algorithm[32] in the randomForest R package. First, the 14 253 products that were assigned to one of the 16 flavour categories were used for RF classification. A fivefold cross-validation was used, for which the data were randomly split into five subsets containing approximately the same number of products and similar distributions of the flavour categories. Next, ingredient information about 80% (4/5 subsets) of the products was used to train a model that predicted the class of the other 20% (1/5) of the products; this was done five times. Additional R settings selected included the number of trees (ntree=2000) and the option to return both the predicted class label and the probabilities for each class. Resulting data were used to evaluate the overall prediction accuracy. For this, we determined how many products were assigned to the correct class according to the RF model (ie, the flavour category with the highest probability). In addition, we determined for how many incorrectly assigned products the correct class received the second highest probability according the RF

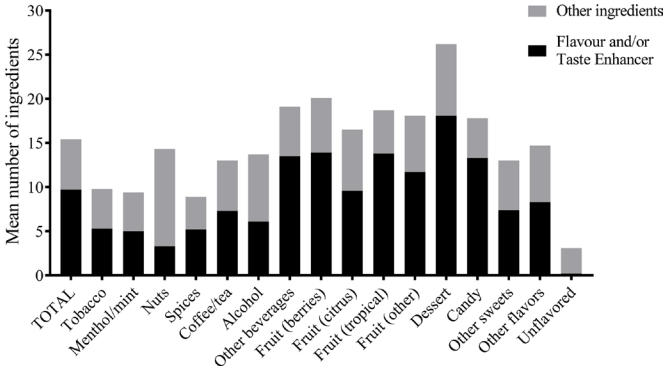

Figure 1 Mean number of ingredients indicated as having a 'flavour and/or taste enhancer' function (black) and ingredients with another function (grey) in total and for each of the separate flavour categories. Other functions of ingredients may include addictiveness enhancers, carriers, casings, fibres, humectants, solvents, processing aids, smoke odour modifiers, water-wetting agents and viscosity modifiers.[36]

model (including tied second place). To determine the chance-based prediction accuracy, we randomly reassigned each product to one of the categories and repeated the machine learning analysis. This resulted in an overall chance accuracy of 10.2%. Finally, we trained a model using quantitative information about the complete set of 14 253 products with an assigned flavour category to predict the class of the 2585 products defined as 'unclassifiable' in our previous study.[7]

Because quantitative information is not always reported, the analyses were repeated using qualitative information about the ingredients only to provide a proof of principle that the method can also be used for qualitative data.

RESULTS

Mean number of flavourings

Over all 16 839 e-liquids, the mean number of reported flavourings per e-liquid was 10±15. Figure 1 shows the mean number of flavourings and other ingredients in total and for each of the separate flavour categories. The mean number of flavourings per flavour category (excluding *unflavoured*) ranged from 3±8 (for *nuts*) to 18±20 (for *dessert*).

On average, 63% of the total number of ingredients within one e-liquid were flavourings. The mean number of flavourings as percentage of the total number of ingredients (excluding *unflavoured*) was highest for e-liquids classified as *candy* (75% were flavourings) and lowest for *nuts* (23% were flavourings). The median concentration of total flavourings per e-liquid was 28.0 mg/10 mL.

Most frequently added flavourings and their quantities

We identified 219 unique ingredients reported to be added to more than 100 e-liquids of the entire dataset. An overview of these ingredients, including their prevalence, is shown in online supplementary table S1. This overview covers 99.4% of all unique ingredients (n=8352) reported. Ingredients other than flavouring ingredients were glycerol, nicotine, propylene glycol, water, ethanol and triacetin. These compounds were present in, respectively, 94%, 88%, 86%, 45%, 23% and 15% of all e-liquids.

Twenty-five flavouring ingredients were added to more than 10% of the overall sample of e-liquids (see table 1). The most frequently used flavourings were vanillin (present in 35.2% of the total set), ethyl maltol (32.0%) and ethyl butyrate (28.4%). The highest median concentration was reported for menthol (18.4 mg/10 mL) and the lowest median concentration was reported for benzaldehyde (0.3 mg/10 mL).

The five flavourings that were most frequently used per separate flavour category are listed in table 2. Online supplementary table S2 shows an overview of all flavourings added to more than 10% of the e-liquids for each flavour category separately. Only two flavourings, ethyl maltol and vanillin, were added to more than 10% of the e-liquids of all flavour categories (except for *unflavoured*). However, 29 flavourings were added to more than 10% of the e-liquids in a single category (excluding those specific to the *other flavours* category). These 'category-specific' flavourings were: β-damascone (for *tobacco*); eucalyptol, menthone and peppermint oil (for *menthol/mint*); 2,5-dimethylpyrazine, 2–3-hexanedione, 4,5-dimethyl-3-hydroxy-2,5-dihydrofuran-2-one, 5-methyl furfural, γ-dodecalactone and triethyl citrate (for *nuts*); anethole and trans-anethole (for *spices*); ethyl heptanoate, ethyl nonanoate, isoamyl alcohol and lactic acid (for *alcohol*); cinnamaldehyde (for *other beverages*); dimethyl sulfide and propionic acid (for *fruit-berries*); orange oil (for *fruit-citrus*); isobutyl acetate and trans-2-hexenal (for *fruit-other*); and 4-methyl-5-thiazole ethanol, anisyl alcohol, benzyl benzoate, γ-hexalactone, methyl-alpha-ionone, methyl-thio-methyl-pyrazine and propenyl guaethol (for *dessert*). See online supplementary table S3 for flavour descriptions of these ingredients.

Predictive value of flavouring composition

Using information about ingredient quantities, 9982 of 14 253 e-liquids were assigned to the correct flavour category. This means that the overall prediction accuracy of the RF algorithm was 70%. For 3740 incorrectly assigned products (26% of total e-liquid sample), the correct class (ie, the manually assigned class) received the second highest probability.

Using the RF model to predict the flavour category of the e-liquids that were defined as 'unclassifiable' in our previous study[7] resulted in: 56% of these e-liquids being classified as *tobacco*; 10% as *fruit-other*; 9% as *fruit-berries*; 7% as *menthol/mint*; 7% as *dessert*; 3% as *alcohol*; 3% as *other sweets*; 2% as *fruit-citrus*; 2% as *fruit-tropical*; 1% as *coffee/tea*; and 1% as *other beverages* (see online supplementary table S4).

Finally, when only using information about the presence of ingredients rather than their amounts (a qualitative rather than quantitative approach), the RF algorithm had an overall prediction accuracy of 66%.

DISCUSSION

This study provides an overview of frequently added ingredients, and their quantities, in all e-liquids available on the Dutch market in 2017. In total, 219 unique ingredients were identified, covering 99.4% of all ingredients reported to be used in our dataset of 16 839 e-liquids. On average, 63% of the total number of ingredients per e-liquid were flavourings and the mean number of flavourings per e-liquid was 10. Common non-flavouring ingredients were glycerol, nicotine, propylene glycol, water, ethanol and triacetin. Vanillin, also one of the mostly used additives in tobacco products[33] for its sweet, creamy, vanilla-like flavour,[34] was the flavouring most frequently added to e-liquids. The overall median concentration was highest for menthol: a compound that is commonly added to tobacco products for its cooling and smoothing effects.[35] The highest mean numbers of flavourings were found for flavour categories that typically contain sweet e-liquids such as *dessert*, *other beverages*, *fruit*

Table 1 The top 25 most frequently added flavouring ingredients in e-liquids from the EU-CEG dataset

	Flavouring ingredient	Prevalence (% of total e-liquids)	Median concentration (mg/10 mL)	Flavour description	GHS code*
1	Vanillin	35.2	7.0	Sweet, powerful, creamy, vanilla-like.	H317 and H319.
2	Ethyl maltol	32.0	5.9	Sweet, fruity-caramellic, cotton candy.	H302.
3	Ethyl butyrate	28.4	3.6	Ethereal, fruity with buttery-pineapple-banana, ripe fruit and juicy notes.	H226.
4	Ethyl acetate	23.2	1.1	Ethereal, sharp, wine-brandy-like.	H225, H319 and H336.
5	Maltol	22.8	1.3	Sweet, fruity, berry, strawberry, caramellic.	H302.
6	Ethyl vanillin	19.4	6.8	Intense, sweet, creamy, vanilla-like.	H302, H315, H319, H335 and H412.
7	Furaneol	19.3	2.0	Fruity, caramelised, roasted, pineapple-strawberry.	H302, H317 and H319.
8	Methyl cyclopentenolone	18.3	2.0	Very strong, caramellic-maple, lovage.	n/a
9	γ-Decalactone	18.2	0.5	Coconut-peach.	n/a
10	Cis-3-hexenol	17.8	1.5	Strong, fresh, green, grassy.	H226 and H319.
11	Isoamyl acetate	16.3	2.3	Sweet, fruity, banana, pear.	H226.
12	Ethyl 2-methyl butyrate	16.0	2.2	Strong, green, fruity, apple with strawberry notes.	H226.
13	Acetic acid	15.7	1.2	Pungent, sour, acid, vinegar.	H226 and H314.
14	Butyric acid	15.0	0.8	Tropical fruity floral, plum, apricot-pear-like.	H314.
15	Linalool	14.5	0.9	Sweet floral-woody with slight citrus notes.	H315, H317 and H319.
16	Benzyl alcohol	14.2	3.3	Faint, sweet, almond fruity, somewhat chemical.	H302 and H332.
17	Ethyl hexanoate	13.6	0.5	Strong, fruity, pineapple, banana with strawberry, pear and tropical notes.	H226.
18	Benzaldehyde	12.4	0.3	Bitter almond oil, sweet cherry.	H302.
19	Menthol	12.1	18.4	String trigeminal cooling sensation with a slight mint note.	H315 and H319.
20	Isoamyl isovalerate	11.5	0.8	Fruity, green-apple, pineapple, tropical, mango, apricot, cognac.	H411.
21	δ-Decalactone	11.2	0.3	Sweet, creamy, milky, peach, nut, buttery.	n/a
22	Hexanoic acid	11.1	0.4	Heavy, fatty, cheesey-sweaty.	H311, H314 and H318.
23	Ethyl propionate	10.9	0.5	Strong, ethereal, fruity, rum-like.	H225.
24	γ-Undecalactone	10.9	0.4	Strong fatty, peach-apricot.	H411 and H412.
25	Hexyl acetate	10.3	1.0	Sweet, fruity, pear-apple, green, banana.	H411.

Prevalence is reported as the number of e-liquids containing the respective flavouring as percentage of the total number of e-liquids. Flavour descriptions were retrieved from a commercial flavour database.[34] GHS codes were retrieved from PubChem (https://pubchem.ncbi.nlm.nih.gov/)
*GHS=Globally Harmonized System of Classification and Labeling of Chemicals; hazard statements can be found at https://pubchem.ncbi.nlm.nih.gov/ghs/.
EU-CEG, European Common Entry Gate.

and *candy*. As our results showed category-specific flavouring patterns, we were able to successfully predict an e-liquid's flavour category based on patterns of flavouring compositions (70% accuracy).

Main applications

Our study provides a comprehensive overview of flavourings added to e-liquids, including their quantities, thereby giving directions to other researchers for selection of target compounds in their future chemical analytical and toxicological studies regarding e-liquid flavourings. For example, based on their research questions, researchers can use our data to select any number of most frequently used flavourings, or flavourings that are unique to particular flavour categories.

In addition, our study showed that using a machine-learning algorithm on a dataset on e-liquid flavouring compositions can provide a reliable estimation of marketed e-liquid flavours. The algorithm can be successfully applied using both ingredient quantities as well as qualitative information only. The resulting overview of marketed e-liquid flavours and their distribution across categories can be used for comparative analyses between regions or countries and to keep track of market trends. Finally, regulators can combine our results with sensory data on flavour preferences from (potential) consumers to define regulatory targets for reducing e-liquid appeal and use among vulnerable user groups (eg, non-smokers and youth).

Considerations regarding the information source

It should be noted that the information used in this study was provided by the industry, without the aim of sustaining research, and were not validated by an independent party. Therefore, while manufacturers were instructed to report ingredient quantities in weight per product unit (ie, mg/10 mL),[36] it is uncertain whether all ingredients were reported in the correct unit. Due to the large size of the dataset used in this study, potential data artefacts are expected to not have influenced overall results. Yet, as information was provided by the industry, results should be interpreted with appropriate caution. In addition, as the EU-CEG information used in this study was extracted at a single time point, results may not be generalisable over time. Therefore, regular follow-up analyses would be worthwhile to get more insight into the dynamics of the e-liquid flavour market.

Benefits and limitations of automatic classification

The overall accuracy of predicting an e-liquid's flavour category based on patterns of flavouring compositions was 70% using ingredient quantities. This is almost seven times higher

Table 2 Overview of the top five most frequently added flavourings for each individual flavour category[30]

Flavour category	Top five flavouring ingredients	Prevalence (%)*	Median concentration (mg/10 mL)
Tobacco	Ethyl maltol	31.2	6.8
	Methyl cyclopentenolone	29.6	1.5
	Vanillin	25.9	4.1
	2,3,5-Trimethylpyrazine	15.8	1.3
	Furaneol	13.1	3.0
Menthol/mint	Menthol	58.6	57.6
	Menthone	16.6	22.7
	Ethyl maltol	12.7	0.7
	Vanillin	11.9	1.2
	Eucalyptol	11.5	7.2
Nuts	Vanillin	58.8	30.6
	Ethyl maltol	47.5	24.5
	Ethyl vanillin	35.0	32.0
	Acetoin	27.5	24.0
	Maltol	26.9	6.2
Spices	Menthol	23.5	20.4
	Anethole	22.8	60.0
	Ethyl maltol	22.8	17.7
	Ethyl butyrate	13.0	2.5
	Benzyl alcohol	11.7	28.0
Coffee/tea	Vanillin	36.0	5.7
	Methyl cyclopentenolone	21.4	5.7
	Benzyl alcohol	20.9	17.0
	Ethyl maltol	15.3	3.3
	Ethyl vanillin	15.0	2.5
Alcohol	Vanillin	50.3	9.0
	Ethyl acetate	25.7	7.9
	Ethyl butyrate	23.1	6.0
	Ethyl propionate	18.8	15.0
	Ethyl heptanoate	18.5	2.2
Other beverages	Vanillin	44.5	7.2
	Ethyl butyrate	39.6	2.2
	Ethyl maltol	36.6	3.2
	Ethyl acetate	30.9	0.8
	Ethyl vanillin	30.8	5.4
Fruit (berries)	Ethyl butyrate	52.4	6.2
	Cis-3-hexenol	41.2	2.3
	Vanillin	39.1	3.1
	Furaneol	36.8	3.3
	Ethyl acetate	36.0	1.4
Fruit (citrus)	Ethyl maltol	32.9	6.0
	Ethyl butyrate	31.1	2.9
	Vanillin	31.0	6.0
	Ethyl acetate	26.3	0.7
	Linalool	25.8	1.6
Fruit (tropical)	Ethyl butyrate	38.5	2.4
	Vanillin	36.5	2.8
	Isoamyl acetate	32.0	2.2
	Ethyl acetate	31.2	1.1
	Ethyl maltol	30.7	1.3
Fruit (other)	Ethyl butyrate	42.6	3.8
	Ethyl acetate	37.7	3.1
	Isoamyl acetate	35.9	3.6
	Vanillin	32.0	2.8
	Maltol	28.1	1.4
Dessert	Vanillin	74.5	21.8
	Ethyl maltol	64.8	9.0
	Ethyl vanillin	63.0	13.8
	Maltol	50.9	1.4
	Methyl cyclopentenolone	49.2	1.2
Candy	Ethyl maltol	39.4	0.5
	Isoamyl acetate	37.4	2.3
	Ethyl butyrate	37.2	1.1
	Vanillin	35.2	1.1
	Ethyl acetate	35.0	0.4
Other sweets	Vanillin	61.4	30.0
	Ethyl maltol	37.5	28.0
	Ethyl vanillin	35.0	19.7
	Maltol	26.0	1.3
	Piperonal	24.4	3.2
Other flavours	Linalool	30.5	3.0
	Ethyl butyrate	25.2	6.0
	Ethyl acetate	23.2	8.0
	Maltol	23.2	4.1
	Ethyl 2-methyl butyrate	22.5	5.9
Unflavoured	NA	NA	NA

For a complete overview, see online supplementary table S1. See online supplementary table S3 for flavour descriptions of these ingredients.
*Prevalence is reported as the number of e-liquids containing the respective flavouring as percentage of the total number of e-liquids within that category.

than chance level (10.2%) and was only slightly lower when only qualitative information was used (66%). This shows that our approach provides a reliable estimation of the distribution of flavour categories on market or country level. It should be noted that the algorithm is not necessarily suitable for predicting flavour categories of individual products. Nevertheless, an important benefit of automatically classifying e-liquids over manual classification is that it significantly limits the time and workload needed to classify a large sample of products.

Other important advantages of classification based on ingredients are that it is insensitive to interpretation of flavour descriptions and allows for classification of products with ambiguous brand names (eg, 'Spaceship', hypothetically) or lacking flavour-related marketing information. In our previous study, a set of 2586 e-liquids could not be classified in any flavour category, due to unavailable or unspecific flavour-related information. However, using the RF algorithm and ingredient information, we were able to classify these e-liquids in addition to the other 14 253 products, thus provide a complete census of e-liquids on the Dutch market.

While the overall prediction accuracy was high, 30% of all e-liquids were misclassified. It is important to note that, in this study, an e-liquid was considered 'misclassified' when the RF algorithm assigned it to a flavour category different from the one it was assigned to in the manual classification approach.[7] The term 'misclassified' is debatable: an e-liquid that was misclassified by the RF algorithm may actually better fit the category to which it was assigned by the algorithm than the category to which it was assigned by manual classification. For example, a hypothetical e-liquid marketed as 'strawberry and

mint' would be manually classified in the *fruit-berries* category, because strawberry was the first flavour description mentioned (see classification guidelines in ref 30). However, this e-liquid may contain more menthol-related than strawberry-related flavouring ingredients, thus being potentially classified in the *menthol/mint* category by the RF algorithm. Whether the correct flavour category is *fruit-berries* or *menthol/mint* is debatable, so this type of considerations increases difficulties regarding e-liquid flavour classification.

In this study, for most of the e-liquids that were misclassified (87.5%), the correct class received the second highest probability. This shows that, if the algorithm assigned a product to the wrong class, the correct flavour category was usually second best. In these cases, the e-liquid flavour closely relates to multiple flavour categories. This can be caused by one e-liquid containing ingredients with flavour descriptions related to multiple categories, or ambiguous flavour-related marketing information that can be interpreted as relating to more than one flavour category. For example, an e-liquid marketed as 'apple pie' (and thus classified in the main category *dessert*) may contain a mixture of ingredients that are separately described as having an apple, cinnamon or pie flavour. Due to this combination of flavourings, this e-liquid may be as much likely to be classified in the *fruit-other, spices* or *dessert* category. Other examples of such misclassifications in our dataset are: (1) an e-liquid with 'peanut butter, vanilla & banana' flavour, incorrectly assigned to *other sweets* while manually classified as *nuts* and (2) an e-liquid with 'lemonade' flavour, incorrectly assigned to *fruit-citrus* while manually classified as *other beverages*. The existence of products with such combined flavours makes both automatic and manual flavour classification challenging.

To improve classification results, classification of some products could be reconsidered, particularly of the products of which the probability of the assigned category is rather low, or of which probabilities of the assigned category and the second-best category differ not much. For example, this could be a product that is classified in the *fruit-other* category with a probability of 49% and received a probability of 46% for the *dessert* category (second-best). These e-liquids could be manually reclassified, for example, using flavour-related information from web shops or using sensory analysis, provided that these e-liquids are still on the market at the time of data analysis. It should be noted that, although classifying e-liquids based on human expert judgement would be an accurate approach, training a panel to identify each e-liquid's flavour would be time-consuming.

Most misclassifications occurred in categories containing a low number of products (ie, *unflavoured, nuts* and *spices*), which may be due to the fact that they contained fewer examples for training the algorithm. Besides, a relatively high number of products from the *other flavours* category were incorrectly assigned to other categories. This seems inevitable as the e-liquids in this category do not have a certain common flavour but instead have miscellaneous flavours that did not fit in any other category. Hence, correct classification for the *other flavours* category is more difficult but also less relevant as this category is by definition a heterogeneous group.

Finally, it should be noted that our analyses were based on flavourings that were reported to be added to more than 100 e-liquids of the entire dataset (ie, more than 0.6% of the 16 839 products). Hence, a few seldom reported ingredients are not represented in our comprehensive overview of e-liquid ingredients. In addition, only e-liquids with a volume of 10 mL were included in this study. Therefore, results do not include cartridges and prefilled e-cigarettes.

CONCLUSIONS AND RECOMMENDATIONS

We analysed e-liquid ingredients and provided an overview of the flavouring ingredients most frequently reported to be added to e-liquids and their quantities. Besides similarities between flavour categories, we identified flavourings that were specific to only one flavour category. Moreover, we successfully predicted e-liquids' flavour categories based on their flavouring composition with 70% accuracy. Automatically classifying e-liquids in this way allows to quickly generate an overview of marketed e-liquid flavour descriptions in a particular country. Thus, we recommended regulators to request information from manufacturers regarding the compositions of all marketed e-liquids to be able to perform similar analyses. This may help to support compliance and control of potential future regulations of e-liquid flavours.

What this paper adds

What is already known on this subject

- Flavouring ingredients increase e-cigarette attractiveness and use, particularly for young non-smokers, and thereby increase exposure to potentially toxic ingredients.
- Chemical analysis can be used to identify and quantify e-liquid flavouring ingredients.

What important gaps in knowledge exist on this topic

- A comprehensive overview of e-liquid ingredients is needed as a starting point for chemical analytical and toxicological research and for regulatory approaches aimed at reducing e-cigarette attractiveness.
- Instead of manual classification of e-liquids by flavour using marketing descriptions, an automatic approach is needed to create market overviews of e-liquid flavours worldwide in a time-efficient manner.

What this paper adds

- This study is the first to use information from manufacturers to create a comprehensive overview of e-liquid flavourings and their quantities, which may give direction to future research on e-liquid flavourings and may support flavouring-level regulation to decrease attractiveness of e-liquids.
- Our approach of automatically classifying e-liquids into flavour categories using their compositions was successful (70% accuracy) and allows to quickly generate market overviews of e-liquid flavours that can be compared between countries.

Contributors EJZK and AH contributed equally. EJZK, AH and RT conceptualised and designed the study. JP prepared and analysed the data. EJZK and AH interpreted the data and wrote the manuscript. SB, KdG, JP and RT provided critical feedback on the manuscript.

Funding The Ministry of Health, Welfare and Sport (VWS), project 5.7.1.

Competing interests None declared.

Patient consent for publication Not required.

Provenance and peer review Not commissioned; externally peer reviewed.

Data availability statement Processed data are available on reasonable request to the first author. Raw data are confidential and cannot be shared.

ORCID iDs
Erna J Z Krüsemann http://orcid.org/0000-0003-3552-8190
Anne Havermans http://orcid.org/0000-0003-2874-6560

REFERENCES

1 Filippidis FT, Laverty AA, Gerovasili V, *et al*. Two-Year trends and predictors of e-cigarette use in 27 European Union member states. *Tob Control* 2017;26:98–104.
2 McMillen RC, Gottlieb MA, Shaefer RMW, *et al*. Trends in electronic cigarette use among U.S. adults: use is increasing in both smokers and nonsmokers. *Nicotine Tob Res* 2015;17:1195–202.
3 Zare S, Nemati M, Zheng Y. A systematic review of consumer preference for e-cigarette attributes: flavor, nicotine strength, and type. *PLoS One* 2018;13:e0194145.
4 Kong G, Morean ME, Cavallo DA, *et al*. Reasons for electronic cigarette experimentation and discontinuation among adolescents and young adults. *Nicotine Tob Res* 2015;17:847–54.
5 Harrell MB, Weaver SR, Loukas A, *et al*. Flavored e-cigarette use: characterizing youth, young adult, and adult users. *Prev Med Rep* 2017;5:33–40.
6 Pepper JK, Ribisl KM, Brewer NT. Adolescents' interest in trying flavoured e-cigarettes. *Tob Control* 2016;25:ii62–6.
7 Havermans A, Krüsemann EJZ, Pennings J, *et al*. Nearly 20 000 e-liquids and 250 unique flavour descriptions: an overview of the Dutch market based on information from manufacturers. *Tob Control* 2021;30:57–62.
8 Schneller LM, Bansal-Travers M, Goniewicz ML, *et al*. Use of flavored electronic cigarette refill liquids among adults and youth in the US-Results from wave 2 of the population assessment of tobacco and health study (2014-2015). *PLoS One* 2018;13:e0202744.
9 Romijnders KAGJ, van Osch L, de Vries H, *et al*. Perceptions and reasons regarding e-cigarette use among users and Non-Users: a narrative literature review. *Int J Environ Res Public Health* 2018;15. doi:10.3390/ijerph15061190. [Epub ahead of print: 06 Jun 2018].
10 Prochaska JJ. The public health consequences of e-cigarettes: a review by the National academies of sciences. A call for more research, a need for regulatory action. *Addiction* 2019;114:587–9.
11 Chen J, Bullen C, Dirks K. A comparative health risk assessment of electronic cigarettes and conventional cigarettes. *Int J Environ Res Public Health* 2017;14:382.
12 Goniewicz ML, Gawron M, Smith DM, *et al*. Exposure to nicotine and selected toxicants in cigarette smokers who switched to electronic cigarettes: a longitudinal Within-Subjects observational study. *Nicotine Tob Res* 2017;19:160–7.
13 Levy DT, Borland R, Lindblom EN, *et al*. Potential deaths averted in USA by replacing cigarettes with e-cigarettes. *Tob Control* 2018;27:18–25.
14 Margham J, McAdam K, Forster M, *et al*. Chemical composition of aerosol from an e-cigarette: a quantitative comparison with cigarette smoke. *Chem Res Toxicol* 2016;29:1662–78.
15 Cullen KA, Ambrose BK, Gentzke AS, *et al*. Notes from the Field: Use of Electronic Cigarettes and Any Tobacco Product Among Middle and High School Students - United States, 2011-2018. *MMWR Morb Mortal Wkly Rep* 2018;67:1276–7.
16 Miech R, Johnston L, O'Malley PM, *et al*. Adolescent Vaping and Nicotine Use in 2017-2018 - U.S. National Estimates. *N Engl J Med* 2019;380:192–3.
17 Central Agency for Statistics (CBS). *National Institute for public health and the environment (RIVM); Trimbos-instituut*. Roken: Cijfers & Context, 2018.
18 McRobbie H. Modelling the population health effects of e-cigarettes use: current data can help guide future policy decisions. *Nicotine Tob Res* 2017;19:131–2.
19 Layden JE, Ghinai I, Pray I, *et al*. Pulmonary illness related to e-cigarette use in Illinois and Wisconsin — preliminary report. *N Engl J Med Overseas Ed* 2019.
20 Eddingsaas N, Pagano T, Cummings C, *et al*. Qualitative analysis of E-Liquid emissions as a function of flavor additives using two aerosol capture methods. *Int J Environ Res Public Health* 2018;15. doi:10.3390/ijerph15020323
21 Klager S, Vallarino J, MacNaughton P, *et al*. Flavoring chemicals and aldehydes in e-cigarette emissions. *Environ Sci Technol* 2017;51:10806–13.
22 Maciej GL. Carcinogens and toxicants in e-cigarettes. *Cancer Prevention Research* 2013;6.
23 Soneji S, Barrington-Trimis JL, Wills TA, *et al*. Association between initial use of e-cigarettes and subsequent cigarette smoking among adolescents and young adults: a systematic review and meta-analysis. *JAMA Pediatr* 2017;171:788–97.
24 Trump administration combating epidemic of youth e-cigarette use with plan to clear market of Unauthorized, Non-Tobacco-Flavored e-cigarette products. *FDA news release* 2019.
25 Girvalaki C, Tzatzarakis M, Kyriakos CN, *et al*. Composition and chemical health hazards of the most common electronic cigarette liquids in nine European countries. *Inhal Toxicol* 2018;30:361–9.
26 Hahn J, Monakhova YB, Hengen J, *et al*. Electronic cigarettes: overview of chemical composition and exposure estimation. *Tob Induc Dis* 2014;12:23.
27 Hutzler C, Paschke M, Kruschinski S, *et al*. Chemical hazards present in liquids and vapors of electronic cigarettes. *Arch Toxicol* 2014;88:1295–308.
28 Lisko JG, Tran H, Stanfill SB, *et al*. Chemical composition and evaluation of nicotine, tobacco alkaloids, pH, and selected flavors in e-cigarette cartridges and refill solutions. *Nicotine Tob Res* 2015;17:1270–8.
29 Tierney PA, Karpinski CD, Brown JE, *et al*. Flavour chemicals in electronic cigarette fluids. *Tob Control* 2016;25:e10–15.
30 Krüsemann EJZ, Boesveldt S, de Graaf K, *et al*. An E-Liquid flavor wheel: a shared vocabulary based on systematically reviewing E-Liquid flavor classifications in literature. *Nicotine Tob Res* 2019;21:1310–9.
31 European Commission. Eu common entry gate (EU-CEG), 2016. Available: https://ec.europa.eu/health/euceg/introduction_en
32 Breiman L, Forests R. Random forests. *Mach Learn* 2001;45:5–32.
33 van Nierop LE, Pennings JLA, Schenk E, *et al*. Analysis of manufacturer's information on tobacco product additive use. *Tob Regul Sci* 2019;5:182–205.
34 Leffingwell & Associates. Flavor-Base 9 - Tobacco Version for Windows XP/Vista/7&8 2013.
35 Ferris Wayne G, Connolly GN, Application CGN. Application, function, and effects of menthol in cigarettes: a survey of tobacco industry documents. *Nicotine Tob Res* 2004;6 Suppl 1:43–54.
36 The European Commission. Commission Implementing Decision (EU) 2015/2183 of 24 November 2015 establishing a common format for the notification of electronic cigarettes and refill containers (notified under document C(2015) 8087). *Official Journal of the European Union* 2015;309:15–27.

A comprehensive overview of common e-liquid ingredients and how they can be used to predict an e-liquid's flavor category

Erna JZ Krüsemann[1,2*], Anne Havermans[1*], Jeroen LA Pennings[1], Kees de Graaf[2], Sanne Boesveldt[2], Reinskje Talhout[1]

** These authors contributed equally*

[1] Centre for Health Protection, National Institute for Public Health and the Environment (RIVM), Antonie van Leeuwenhoeklaan 9, 3721 MA Bilthoven, The Netherlands

[2] Division of Human Nutrition and Health, Wageningen University, Stippeneng 4, 6708 WE Wageningen, The Netherlands

Corresponding author: Erna JZ Krüsemann, Centre for Health Protection, National Institute for Public Health and the Environment (RIVM), Antonie van Leeuwenhoeklaan 9, 3721 MA Bilthoven, The Netherlands, erna.krusemann@rivm.nl, +31 302742512

Table of Contents

Krüsemann EJZ, *et al. Tob Control* 2020;0:1–7. doi: 10.1136/tobaccocontrol-2019-055447

Supplementary Table S1

Table S1: An overview of all 219 unique ingredients reported to be added to at least 100 e-liquids of the entire data set, including the percentage of e-liquids containing the ingredient. This overview covers 99.4% of all ingredients reported. The non-flavoring ingredients (n=6) are marked in grey and highlighted with an asterisk (*).

E-liquid ingredient	E-liquids containing the ingredient (% of total number of e-liquids)
2,3-Pentanedione	3.09%
1-Pentanol	1.05%
1,4-Dimethoxybenzene	1.91%
2-3-Hexanedione	1.07%
2-Acetyl Furan	1.18%
2-Acetyl Pyridine	1.96%
2-Acetylpyrazine	6.00%
2-Ethyl-3-methyl pyrazine	0.79%
2-Isopropyl-4-methyl thiazole	0.47%
2-Isopropyl-N,2,3-trimethylbutyramide	1.28%
2-Methoxy-3-methyl pyrazine	0.47%
2-Methyl Butyric Acid	9.77%
2-Methylbutyl acetate	2.23%
2-Phenylethanol	4.20%
2-Phenylethyl Acetate	0.57%
2-Propanol	1.43%
2,3-Dimethylpyrazine	2.08%
2,3,5-Trimethylpyrazine	5.67%
2,3,5,6-Tetramethylpyrazine	2.90%
2,5-Dimethylpyrazine	1.54%
2,6-Dimethyl-5-heptenal	2.15%
4-(4-Hydroxyphenyl)-2-butanone	9.43%
4-(4-methoxyphenyl)butan-2-one	1.06%
4-Methyl-5-Thiazole Ethanol	5.53%
4-Terpinenol	0.37%
4,5-Dimethyl-3-Hydroxy-2,5-Dihydrofuran-2-One	1.38%
5-methyl-2-Phenyl-2-Hexenal	1.10%
5-Methyl Furfural	2.73%
6-Methyl-5-Hepten-2-one	0.80%
6-Methyl Coumarin	0.60%
Acetal	1.08%
Acetaldehyde	2.02%
Acetic Acid	15.67%
Acetoin	7.47%
Allyl Hexanoate	5.54%
alpha-Damascenone	0.02%
alpha-Damascone	1.84%
Alpha-Ionone	5.55%
alpha-Methylbenzyl acetate	1.80%
alpha-Pinene	2.14%
alpha-Terpineol	5.51%

Krüsemann EJZ, *et al. Tob Control* 2020;0:1–7. doi: 10.1136/tobaccocontrol-2019-055447

Amyl acetate	1.88%
Amyl Butyrate	1.10%
Anethole	2.30%
Anisaldehyde	9.01%
Anise Oil	0.97%
Anisyl Acetate	0.83%
Anisyl Alcohol	7.04%
Benzaldehyde	12.43%
Benzaldehyde propylene glycol acetal	0.74%
Benzyl Acetate	9.20%
Benzyl Alcohol	14.19%
Benzyl Benzoate	3.10%
Benzyl Butyrate	1.41%
beta-Caryophyllene	1.81%
beta-Damascenone	4.42%
beta-Damascone	5.53%
Beta-Ionone	7.54%
beta-Pinene	2.14%
Blood Orange Oil	1.48%
Bucchu Leaf Oil	1.28%
Butyl Acetate	2.33%
Butyl Butyrate	2.82%
Butyl Butyryl Lactate	3.95%
Butyric Acid	14.99%
Capsicum oleoresin	0.48%
Caramel	2.38%
Carob	1.47%
Carvone	1.45%
Cassia oil	1.25%
Cedrol	2.08%
Cinnamaldehyde	3.78%
Cinnamon Oil	0.99%
Cinnamyl alcohol	0.82%
Cis-3-hexenol	17.77%
Cis-3-hexenyl Acetate	9.18%
Cis-3-Hexenyl Butyrate	3.35%
Cis-3-Hexenyl Isovalerate	0.78%
cis-6-Nonen-1-ol	0.93%
Citral	5.62%
Citric Acid	3.10%
Citronellol	1.88%
Citronellyl Acetate	1.50%
Cocoa Extract	1.54%
Coffee Extract	0.34%
Cornmint Oil	1.26%
D-Carvone	0.35%
Decanal	1.25%
Decanoic Acid	1.94%
delta-Decalactone	11.21%
delta-Dodecalactone	8.65%

Dihydrocoumarin	5.89%
Dimethyl Anthranilate	1.32%
Dimethyl Sulfide	5.02%
Dodecane	0.74%
Ethanol*	23.12%
Ethyl-3-Hydroxy Butyrate	0.48%
Ethyl 2-Methyl Butyrate	15.99%
Ethyl 2-Phenyl Acetate	1.41%
Ethyl Acetate	23.23%
Ethyl Acetoacetate	3.27%
Ethyl Butyrate	28.37%
Ethyl Cinnamate	1.50%
Ethyl Decanoate	1.63%
Ethyl Dodecanoate	0.84%
Ethyl Heptanoate	2.00%
Ethyl Hexanoate	13.56%
Ethyl Isovalerate	9.51%
Ethyl Lactate	3.60%
Ethyl maltol	32.01%
Ethyl menthane carboxamide	1.51%
Ethyl methyl phenylglycidate	2.41%
Ethyl Nonanoate	1.50%
Ethyl Octanoate	2.26%
Ethyl Propionate	10.94%
Ethyl Vanillin	19.38%
Ethyl Vanillin Propylene Glycol Acetal	0.32%
Eucalyptol	3.00%
Eugenol	6.01%
Fenugreek	1.62%
Furaneol	19.31%
Furfural	2.98%
Furfuryl Alcohol	2.32%
gamma-Decalactone	18.16%
gamma-Dodecalactone	2.99%
gamma-Hexalactone	5.10%
gamma-Nonalactone	9.53%
gamma-Octalactone	7.27%
gamma-Terpinene	0.82%
gamma-Undecalactone	10.86%
gamma-Valerolactone	0.74%
Geranial	0.80%
Geraniol	4.06%
Geranyl acetate	3.51%
Ginger oil	0.14%
Glycerol*	94.14%
Glyceryl 1-acetate	0.91%
Guaiacol	6.05%
Hexanal	3.63%
Hexanoic Acid	11.10%
Hexyl Acetate	10.35%

Hexyl Butyrate	1.71%
Ionone (mixed isomers)	0.99%
Isoamyl Acetate	16.33%
Isoamyl Alcohol	4.49%
Isoamyl Butyrate	6.83%
Isoamyl Isovalerate	11.46%
Isoamyl Phenyl Acetate	1.49%
Isobutyl Acetate	4.88%
Isobutyl Alcohol	1.41%
Isobutyl Butyrate	0.69%
Isobutyric Acid	0.50%
Isovaleraldehyde	1.78%
Jasmine Absolute	0.89%
L-Carvone	0.80%
Lactic Acid	3.42%
Lemon oil	6.35%
Levulinic Acid	0.81%
Lime oil	3.98%
Limonene	7.90%
Linalool	14.53%
Linalyl Acetate	3.41%
Maltol	22.76%
Mandarin Oil	0.50%
Menthol	12.10%
Menthone	2.92%
Menthyl acetate	2.32%
Methyl-alpha-ionone	2.62%
Methyl Anthranilate	3.74%
Methyl Cinnamate	9.37%
Methyl cyclopentenolone	18.32%
Methyl dihydrojasmonate	0.81%
Methyl Salicylate	1.19%
Methyl Thiobutyrate	1.47%
Methyl trans-cinnamate	0.77%
Methyl-thio-methylpyrazine	2.45%
Myrcene	1.85%
n-Butanol	1.25%
n-Hexanol	3.58%
n-Octanal	1.10%
n-Propanol	0.51%
Neral	0.75%
Nerol	1.24%
Neryl Acetate	1.59%
Nicotine*	88.41%
Octanoic Acid	4.39%
Octanol	0.74%
Oleic Acid	1.09%
Orange oil	3.72%
p-Cymene	0.96%
Peppermint Oil	1.78%

Peru Balsam	1.16%
Piperonal	9.56%
Potassium Sorbate	1.12%
Propenyl Guaethol	2.42%
Propionic Acid	4.27%
Propyl Acetate	0.33%
Propylene Glycol*	85.78%
Sodium Benzoate	1.16%
Sodium Citrate	1.16%
Spearmint oil	1.07%
Strawberry Extract	1.59%
Sucralose	8.25%
Sugar	1.26%
Tabanone	0.53%
Terpinolene	0.84%
Thio Menthone	1.62%
trans-2-Hexenal	3.85%
trans-2-Hexenoic acid	1.31%
trans-2-Hexenol	1.82%
trans-2-Hexenylacetaat	0.94%
trans-Anethole	2.19%
Triacetin*	14.44%
Triethyl Citrate	4.10%
Vanilla Extract	0.79%
Vanillin	35.17%
Vanillin Propylene Glycol Acetal	1.33%
Veratraldehyde	1.70%
Water*	44.96%

Supplementary Table S2

Table S2: An overview of all flavorings added to more than 10% of the e-liquids within a flavor category [1]. Prevalence is reported as the number of e-liquids containing the respective flavoring as percentage of the total number of e-liquids. The final column shows the median concentration of the flavoring ingredient for the respective flavor category. The 29 "category-specific" flavorings that were added to more than 10% of the e-liquids in a single category are marked in grey.

Category	Flavorings added to >10% of the e-liquid within the respective flavor category	Prevalence (% of total e-liquids)	Median concentration (mg/10mL)
Tobacco *(8 flavorings)*	Ethyl maltol	31.2%	6.78
	Methyl cyclopentenolone	29.6%	1.50
	Vanillin	25.9%	4.11
	2,3,5-Trimethylpyrazine	15.8%	1.32
	Furaneol	13.1%	2.97
	β-Damascone	12.6%	1.10
	2-Acetylpyrazine	12.3%	1.32
	Benzyl Alcohol	10.3%	6.00
Menthol/mint *(6 flavorings)*	Menthol	58.6%	57.60
	Menthone	16.6%	22.68
	Ethyl maltol	12.7%	0.74
	Vanillin	11.9%	1.24
	Eucalyptol	11.5%	7.15
	Peppermint Oil	11.4%	9.87
Nuts *(32 flavorings)*	Vanillin	58.8%	30.63
	Ethyl maltol	47.5%	24.50
	Ethyl Vanillin	35.0%	32.00
	Acetoin	27.5%	24.00
	Maltol	26.9%	6.17
	2,3-Pentanedione	23.8%	24.00
	Methyl cyclopentenolone	23.8%	24.00
	Butyric Acid	22.5%	24.00
	δ-Decalactone	22.5%	6.43
	γ-Decalactone	22.5%	1.63
	Ethyl Acetate	21.9%	24.00
	Ethyl Butyrate	21.9%	24.00
	Acetic Acid	21.3%	24.00
	Furaneol	20.0%	28.56
	Anisaldehyde	19.4%	24.00
	Guaiacol	19.4%	24.00
	δ-Dodecalactone	17.5%	1.63
	2,3,5-Trimethylpyrazine	16.9%	1.10
	Ethyl Propionate	16.9%	24.00
	2-Acetylpyrazine	15.6%	10.48
	Butyl Butyryl Lactate	15.6%	50.91
	γ-Octalactone	15.0%	35.14
	2,5-Dimethylpyrazine	13.8%	0.56
	4,5-Dimethyl-3-Hydroxy-2,5-	11.9%	3.26

	Dihydrofuran-2-One		
	5-Methyl Furfural	11.9%	4.63
	Benzyl Alcohol	11.9%	8.00
	Ethyl Lactate	11.9%	46.28
	γ-Dodecalactone	11.9%	1.63
	Triethyl Citrate	11.9%	46.28
	Piperonal	10.6%	9.00
	Veratraldehyde	10.6%	2.17
	2-3-Hexanedione	10.0%	1.63
Spices *(9 flavorings)*	Menthol	23.5%	20.42
	Anethole	22.8%	60.00
	Ethyl maltol	22.8%	17.72
	Ethyl Butyrate	13.0%	2.49
	Benzyl Alcohol	11.7%	28.00
	Ethyl Vanillin	11.7%	27.37
	trans-Anethole	11.1%	105.00
	Vanillin	11.1%	30.60
	Anisaldehyde	10.5%	0.36
Coffee/tea *(13 flavorings)*	Vanillin	36.0%	5.68
	Methyl cyclopentenolone	21.4%	5.70
	Benzyl Alcohol	20.9%	17.00
	Ethyl maltol	15.3%	3.25
	Ethyl Vanillin	15.0%	2.47
	Furaneol	15.0%	2.35
	Maltol	15.0%	3.00
	δ-Decalactone	14.3%	0.60
	Acetic Acid	13.3%	0.66
	δ-Dodecalactone	11.6%	0.55
	Ethyl Acetate	11.6%	0.47
	Dihydrocoumarin	11.1%	0.97
	Ethyl Butyrate	11.1%	1.28
Alcohol *(22 flavorings)*	Vanillin	50.3%	9.00
	Ethyl Acetate	25.7%	7.92
	Ethyl Butyrate	23.1%	5.98
	Ethyl Propionate	18.8%	15.00
	Ethyl Heptanoate	18.5%	2.23
	Ethyl maltol	17.6%	10.85
	Benzaldehyde	16.5%	6.00
	γ-Nonalactone	15.6%	5.62
	Ethyl Hexanoate	15.3%	0.90
	Acetic Acid	15.0%	1.58
	Maltol	14.2%	4.90
	Allyl Hexanoate	13.9%	2.33
	Piperonal	13.0%	1.14
	δ-Decalactone	12.7%	0.23
	Methyl cyclopentenolone	12.1%	10.82
	Isoamyl Alcohol	11.8%	0.61
	Isoamyl Acetate	11.3%	1.00
	Lactic Acid	11.3%	15.00
	Ethyl Vanillin	11.0%	9.41

	Ethyl Nonanoate	10.4%	1.00
	Furaneol	10.4%	7.96
	Menthol	10.1%	11.75
Other beverages *(44 flavorings)*	Vanillin	44.5%	7.19
	Ethyl Butyrate	39.6%	2.20
	Ethyl maltol	36.6%	3.23
	Ethyl Acetate	30.9%	0.75
	Ethyl Vanillin	30.8%	5.39
	Maltol	28.2%	0.90
	Isoamyl Acetate	26.4%	1.75
	Ethyl 2-Methyl Butyrate	25.5%	0.25
	Furaneol	25.0%	0.66
	Ethyl Hexanoate	23.7%	0.37
	Linalool	23.5%	0.31
	Methyl Cinnamate	22.0%	0.18
	Cis-3-hexenol	21.9%	1.00
	γ-Decalactone	20.6%	0.25
	Butyric Acid	20.3%	0.26
	Ethyl Isovalerate	20.2%	0.90
	Acetic Acid	19.4%	0.14
	Benzyl Acetate	18.5%	1.00
	Isoamyl Isovalerate	17.6%	0.20
	Benzaldehyde	17.2%	0.19
	Lemon oil	17.2%	6.10
	Benzyl Alcohol	16.0%	2.82
	Sucralose	16.0%	8.84
	Citral	15.1%	0.61
	Methyl cyclopentenolone	14.5%	0.46
	Isoamyl Butyrate	14.2%	1.64
	β-Ionone	13.8%	0.75
	Eugenol	13.8%	1.97
	Ethyl Propionate	13.6%	0.48
	Hexanoic Acid	13.6%	0.26
	γ-Undecalactone	13.5%	0.41
	4-(4-Hydroxyphenyl)-2-butanone	13.2%	3.11
	Lime oil	13.1%	27.50
	Alpha-Ionone	12.9%	1.00
	alpha-Terpineol	12.9%	0.74
	Cinnamaldehyde	12.5%	3.31
	Limonene	12.5%	4.50
	Cis-3-hexenyl Acetate	12.3%	0.14
	Piperonal	11.9%	0.20
	Hexyl Acetate	11.4%	0.60
	Allyl Hexanoate	10.9%	1.00
	Anisaldehyde	10.6%	0.11
	γ-Nonalactone	10.3%	0.76
	2-Methyl Butyric Acid	10.2%	0.17
Fruit (berries) *(36 flavorings)*	Ethyl Butyrate	52.4%	6.21
	Cis-3-hexenol	41.2%	2.33
	Vanillin	39.1%	3.05

	Furaneol	36.8%	3.26
	Ethyl Acetate	36.0%	1.36
	Ethyl 2-Methyl Butyrate	35.7%	2.51
	Acetic Acid	33.5%	2.31
	Ethyl maltol	33.5%	3.34
	γ-Decalactone	33.1%	0.91
	Maltol	32.7%	1.21
	4-(4-Hydroxyphenyl)-2-butanone	32.6%	1.78
	Linalool	31.1%	0.92
	Hexanoic Acid	26.3%	0.60
	2-Methyl Butyric Acid	25.7%	2.79
	Butyric Acid	25.5%	0.85
	Ethyl Hexanoate	24.2%	0.81
	Ethyl Isovalerate	23.8%	0.65
	Methyl Cinnamate	23.4%	0.92
	Isoamyl Acetate	22.0%	1.00
	Cis-3-hexenyl Acetate	21.5%	1.00
	Benzyl Acetate	19.6%	0.74
	γ-Undecalactone	19.1%	0.21
	Benzyl Alcohol	19.0%	1.02
	Hexyl Acetate	18.7%	0.41
	Ethyl Vanillin	18.6%	2.92
	β-Ionone	17.4%	0.36
	Ethyl Propionate	17.1%	0.32
	Benzaldehyde	16.1%	0.15
	Alpha-Ionone	16.0%	0.40
	Menthol	15.7%	9.12
	Dimethyl Sulfide	14.1%	0.20
	Isoamyl Isovalerate	13.1%	0.96
	Isoamyl Butyrate	12.8%	0.36
	δ-Decalactone	11.6%	0.13
	Sucralose	11.3%	12.15
	Propionic Acid	11.0%	3.00
Fruit (citrus) *(31 flavorings)*	Ethyl maltol	32.9%	5.98
	Ethyl Butyrate	31.1%	2.85
	Vanillin	31.0%	6.00
	Ethyl Acetate	26.3%	0.66
	Linalool	25.8%	1.56
	Citral	23.4%	2.81
	Lemon oil	21.9%	12.00
	Maltol	20.3%	0.48
	Ethyl Vanillin	19.9%	16.00
	Lime oil	19.4%	9.66
	alpha-Terpineol	19.0%	0.26
	Furaneol	18.4%	1.50
	Cis-3-hexenol	17.6%	0.77
	Limonene	17.1%	13.00
	Orange oil	16.1%	1.74
	γ-Decalactone	15.9%	0.24
	Benzyl Alcohol	15.3%	1.20

	Methyl cyclopentenolone	14.9%	8.57
	Ethyl Hexanoate	14.7%	2.62
	δ-Dodecalactone	14.1%	1.13
	Ethyl 2-Methyl Butyrate	13.1%	0.75
	Acetic Acid	12.7%	1.10
	Sucralose	12.4%	13.88
	γ-Nonalactone	12.1%	2.05
	Ethyl Propionate	11.1%	1.30
	4-(4-Hydroxyphenyl)-2-butanone	10.6%	1.00
	Butyric Acid	10.3%	0.10
	Hexanoic Acid	10.3%	0.68
	Isoamyl Isovalerate	10.3%	0.67
	2-Methyl Butyric Acid	10.2%	2.64
	Ethyl Isovalerate	10.0%	0.20
Fruit (tropical) *(39 flavorings)*	Ethyl Butyrate	38.5%	2.42
	Vanillin	36.5%	2.75
	Isoamyl Acetate	32.0%	2.25
	Ethyl Acetate	31.2%	1.09
	Ethyl maltol	30.7%	1.31
	Cis-3-hexenol	29.1%	1.00
	Ethyl Hexanoate	27.2%	0.72
	Furaneol	26.2%	0.72
	Maltol	24.9%	1.04
	γ-Decalactone	23.8%	0.49
	Allyl Hexanoate	22.8%	1.11
	Acetic Acid	21.8%	1.20
	Ethyl 2-Methyl Butyrate	19.7%	1.47
	Ethyl Vanillin	19.6%	1.78
	γ-Undecalactone	19.2%	0.33
	Isoamyl Isovalerate	19.2%	1.10
	Butyric Acid	18.7%	1.02
	Benzyl Alcohol	18.4%	2.73
	γ-Nonalactone	18.4%	0.78
	Benzaldehyde	17.6%	0.20
	Isoamyl Butyrate	15.8%	1.54
	Benzyl Acetate	15.4%	0.24
	Methyl cyclopentenolone	15.4%	0.33
	Limonene	15.3%	2.99
	Hexanoic Acid	15.2%	0.40
	Ethyl Propionate	15.0%	0.37
	γ-Octalactone	14.7%	0.42
	Linalool	14.7%	0.40
	Ethyl Isovalerate	14.4%	0.25
	2-Methyl Butyric Acid	13.4%	1.40
	Lemon oil	13.4%	0.25
	Cis-3-hexenyl Acetate	12.7%	1.00
	Hexyl Acetate	12.7%	0.83
	Methyl Cinnamate	12.2%	0.20
	Sucralose	11.8%	8.98
	δ-Decalactone	11.4%	0.24

	Eugenol	11.2%	2.63
	Piperonal	11.1%	0.44
	Acetoin	10.5%	1.00
Fruit (other) *(35 flavorings)*	Ethyl Butyrate	42.6%	3.78
	Ethyl Acetate	37.7%	3.08
	Isoamyl Acetate	35.9%	3.57
	Vanillin	32.0%	2.81
	Maltol	28.1%	1.41
	Ethyl maltol	28.0%	0.95
	Ethyl 2-Methyl Butyrate	25.8%	1.49
	Hexyl Acetate	25.8%	1.16
	Cis-3-hexenol	25.7%	1.17
	Benzaldehyde	24.1%	1.45
	γ-Decalactone	22.4%	0.29
	Acetic Acid	20.5%	1.21
	Linalool	19.9%	0.88
	γ-Undecalactone	19.8%	0.77
	Furaneol	18.6%	1.00
	Ethyl Hexanoate	17.9%	0.24
	Cis-3-hexenyl Acetate	17.7%	0.51
	Benzyl Alcohol	15.3%	1.40
	Benzyl Acetate	14.6%	1.00
	Ethyl Vanillin	14.3%	1.88
	Ethyl Isovalerate	13.9%	1.20
	Isobutyl Acetate	13.7%	1.76
	Butyric Acid	13.2%	0.32
	trans-2-Hexenal	12.9%	1.93
	2-Methyl Butyric Acid	12.0%	1.34
	Ethyl Propionate	11.9%	0.33
	4-(4-Hydroxyphenyl)-2-butanone	11.7%	1.43
	Isoamyl Isovalerate	11.7%	0.77
	δ-Decalactone	11.0%	0.20
	Methyl cyclopentenolone	10.7%	0.50
	β-Ionone	10.5%	0.13
	Limonene	10.5%	1.77
	Anisaldehyde	10.3%	0.92
	Menthol	10.3%	16.59
	Sucralose	10.2%	10.80
Dessert *(47 flavorings)*	Vanillin	74.5%	21.80
	Ethyl maltol	64.8%	9.02
	Ethyl Vanillin	63.0%	13.76
	Maltol	50.9%	1.37
	Methyl cyclopentenolone	49.2%	1.20
	Butyric Acid	45.7%	1.00
	Ethyl Butyrate	44.4%	1.40
	γ-Decalactone	40.6%	0.24
	δ-Decalactone	39.5%	0.36
	Ethyl Acetate	38.7%	0.39
	Furaneol	34.2%	2.10
	γ-Nonalactone	34.1%	0.96

	Isoamyl Isovalerate	33.5%	0.51
	Piperonal	33.1%	0.36
	Anisaldehyde	31.8%	0.20
	4-Methyl-5-Thiazole Ethanol	30.2%	0.35
	Guaiacol	30.0%	0.24
	Ethyl Propionate	29.9%	0.98
	Anisyl Alcohol	29.4%	0.72
	δ-Dodecalactone	29.0%	0.29
	Benzyl Alcohol	28.8%	4.01
	Hexanoic Acid	27.4%	1.14
	γ-Octalactone	27.0%	0.54
	Cis-3-hexenol	25.0%	1.00
	γ-Hexalactone	24.8%	0.54
	Acetoin	24.0%	1.36
	Ethyl 2-Methyl Butyrate	22.2%	1.51
	Sucralose	21.8%	8.68
	Acetic Acid	21.3%	0.44
	Dihydrocoumarin	21.3%	1.00
	2-Acetylpyrazine	20.5%	1.26
	Benzaldehyde	20.0%	0.10
	Ethyl Hexanoate	17.6%	0.62
	Linalool	17.2%	0.79
	2-Methyl Butyric Acid	16.4%	1.18
	Methyl Cinnamate	14.7%	1.00
	Butyl Butyryl Lactate	13.7%	1.01
	2,3-Pentanedione	13.3%	1.17
	γ-Undecalactone	13.3%	0.67
	Benzyl Benzoate	13.1%	0.46
	Isoamyl Acetate	13.1%	2.20
	Methyl-alpha-ionone	12.8%	1.77
	Ethyl Lactate	12.6%	0.67
	Methyl-thio-methylpyrazine	11.9%	0.06
	β-Ionone	11.1%	0.42
	Hexyl Acetate	10.8%	0.60
	Propenyl Guaethol	10.2%	1.00
Candy *(34 flavorings)*	Ethyl maltol	39.4%	0.50
	Isoamyl Acetate	37.4%	2.29
	Ethyl Butyrate	37.2%	1.13
	Vanillin	35.2%	1.07
	Ethyl Acetate	35.0%	0.39
	Maltol	26.2%	0.50
	Cis-3-hexenol	23.2%	1.00
	Ethyl 2-Methyl Butyrate	22.9%	0.17
	Ethyl Vanillin	21.6%	0.28
	Linalool	21.0%	0.66
	Butyric Acid	20.5%	0.32
	Ethyl Hexanoate	20.1%	0.11
	γ-Decalactone	19.9%	0.24
	Ethyl Isovalerate	19.5%	0.17
	Benzaldehyde	17.9%	0.10

	Furaneol	17.6%	1.21
	Benzyl Acetate	17.3%	1.00
	4-(4-Hydroxyphenyl)-2-butanone	17.0%	1.00
	Benzyl Alcohol	16.7%	0.68
	γ-Undecalactone	16.0%	0.20
	Hexyl Acetate	16.0%	1.00
	Isoamyl Butyrate	15.7%	1.28
	Isoamyl Isovalerate	14.8%	0.39
	Methyl Cinnamate	13.9%	0.18
	Methyl cyclopentenolone	13.3%	0.14
	Hexanoic Acid	12.9%	0.22
	2-Methyl Butyric Acid	12.6%	0.30
	Citral	12.3%	0.75
	Acetic Acid	12.2%	0.40
	β-Ionone	12.0%	0.66
	Cis-3-hexenyl Acetate	11.4%	0.26
	Ethyl Propionate	10.6%	0.20
	Eugenol	10.6%	0.45
	Lemon oil	10.4%	0.13
Other sweets *(20 flavorings)*	Vanillin	61.4%	30.00
	Ethyl maltol	37.5%	28.04
	Ethyl Vanillin	35.0%	19.68
	Maltol	26.0%	1.30
	Piperonal	24.4%	3.18
	Methyl cyclopentenolone	22.0%	5.65
	δ-Decalactone	18.9%	0.71
	Ethyl Butyrate	18.6%	0.51
	γ-Nonalactone	16.2%	0.75
	Furaneol	15.1%	2.22
	Isoamyl Isovalerate	13.7%	1.11
	2,3,5-Trimethylpyrazine	13.1%	0.25
	Anisaldehyde	13.1%	0.50
	Ethyl Acetate	13.0%	0.57
	γ-Decalactone	12.9%	0.30
	Butyric Acid	12.6%	0.77
	Acetoin	12.0%	1.15
	Guaiacol	11.1%	0.04
	δ-Dodecalactone	10.0%	0.40
	Veratraldehyde	10.0%	2.96
Other flavors *(34 flavorings)*	Linalool	30.5%	3.00
	Ethyl Butyrate	25.2%	6.00
	Ethyl Acetate	23.2%	8.00
	Maltol	23.2%	4.10
	Ethyl 2-Methyl Butyrate	22.5%	5.90
	Vanillin	21.9%	21.00
	Acetic Acid	21.2%	0.70
	Isoamyl Acetate	20.5%	6.00
	Cis-3-hexenyl Acetate	19.2%	0.30
	Cis-3-hexenol	18.5%	1.80
	Ethyl Acetoacetate	18.5%	0.20

	Ethyl Hexanoate	18.5%	0.20
	Butyric Acid	17.2%	1.95
	Hexyl Acetate	17.2%	8.00
	Isoamyl Butyrate	17.2%	5.20
	Cis-3-Hexenyl Butyrate	15.2%	0.10
	Ethyl Isovalerate	15.2%	0.10
	γ-Decalactone	15.2%	1.20
	Hexanal	15.2%	0.02
	Linalyl Acetate	15.2%	0.04
	Furaneol	13.9%	7.40
	Myrcene	13.2%	25.60
	Anisaldehyde	12.6%	1.80
	Dihydrocoumarin	12.6%	30.00
	Ethyl maltol	12.6%	17.70
	Limonene	12.6%	6.97
	2-Phenylethanol	11.9%	0.10
	Citronellol	11.9%	6.00
	2-Methyl Butyric Acid	11.3%	1.10
	Benzaldehyde	11.3%	6.60
	β-Caryophyllene	11.3%	25.60
	Hexanoic Acid	11.3%	0.20
	Methyl Cinnamate	11.3%	0.40
	β-Damascenone	10.6%	0.10
Unflavored *(0 flavorings)*	NA	NA	NA

Supplementary Table S3

Table S3: Flavor descriptions, retrieved from a commercial flavor database [2], of the ingredients listed in the main text of the manuscript. Ingredients are presented in alphabetical order.

Flavoring ingredient	**Flavor description**	***Flavor category that the ingredient is specific to (category-specific flavorings)****
2,3,5-Trimethylpyrazine	Baked potato, roasted nut, cocoa, coffee, burnt	
2,5-Dimethylpyrazine	Chocolate, roasted nuts, earthy	*Nuts*
2-3-Hexanedione	Creamy, sweet buttery odor & butter-cheese, fruity, caramellic taste	*Nuts*
4,5-Dimethyl-3-hydroxy-2,5-dihydrofuran-2-one	Powerful caramel aroma; sweet burnt taste	*Nuts*
4-Methyl-5-thiazole ethanol	In dilution, meaty-beef like with nutty note	*Dessert*
5-Methyl furfural	Sweet spicy, bready, nutty, caramellic	*Nuts*
Acetic acid^{o}	Pungent, sour, acid, vinegar	
Acetoin	Creamy-buttery, yogurt-like	
Anethole	Sweet, herbaceous, anise (artificial licorice)	*Spices*
Anisyl alcohol	Sweet, fruity, floral, balsamic anisic-vanilla-creamy-coumarinic like	*Dessert*
Benzaldehyde^{o}	Bitter almond oil, sweet cherry	
Benzyl alcohol^{o}	Faint, sweet, almond fruity, somewhat chemical	
Benzyl benzoate	Faint, sweet, balsamic with slight bitter, fruity notes	*Dessert*
Butyric acid^{o}	Tropical fruity floral, plum, apricot-pear-like	
Cinnamaldehyde	Spicy, cinnamon-cassia-like with sweet warm (hot) taste	*Other beverages*
Cis-3-hexenol^{o}	Strong, fresh, green, grassy	
δ-Decalactone^{o}	Sweet, creamy, milky, peach, nut, buttery	
Dimethyl sulfide	Pungent, cabbage, cooked vegetable odor; corn-like on dilution	*Fruit, berries*
Ethyl 2-methyl butyrate^{o}	Strong, green, fruity, apple with strawberry notes	
Ethyl acetate^{o}	Ethereal, sharp, wine-brandy-like	
Ethyl butyrate^{o}	Ethereal, fruity with buttery-pineapple-banana, ripe fruit & juicy notes	
Ethyl heptanoate	Strong, fruity, winey, cognac-like	*Alcohol*
Ethyl hexanoate^{o}	Strong, fruity, pineapple, banana with strawberry, pear & tropical notes	
Ethyl maltol^{o}	Sweet, fruity-caramellic, cotton candy	
Ethyl nonanoate	Fatty-waxy, oily, wine-cognac, grape, tropical, nut-like	*Alcohol*
Ethyl propionate^{o}	Strong, ethereal, fruity, rum-like	
Ethyl vanillin^{o}	Intense, sweet, creamy, vanilla-like	
Eucalyptol	Strong, camphoraceous, cool, fresh	*Menthol/mint*
Furaneol^{o}	Fruity, caramelized, roasted, pineapple-strawberry	
Hexanoic acid^{o}	Heavy, fatty, cheesey-sweaty	
Hexyl acetate^{o}	Sweet, fruity, pear-apple, green, banana	
Isoamyl acetate^{o}	Sweet, fruity, banana, pear	
Isoamyl alcohol	Breathtaking, alcoholic odor; in dilution a winey-brandy taste	*Alcohol*
Isoamyl isovalerate^{o}	Fruity, green-apple, pineapple, tropical, mango, apricot, cognac	
Isobutyl acetate	Fruity, banana-apple-pear-pineapple	*Fruit, other*
Lactic acid	Weak, sour, buttermilk	*Alcohol*
Linalool^{o}	Sweet floral-woody with slight citrus notes	
Maltol^{o}	Sweet, fruity, berry, strawberry, caramellic	
Menthol^{o}	String trigeminal cooling sensation with a slight mint note	

Menthone	Minty-herbaceous (not green); dry woody notes	*Menthol/mint*
Methyl cyclopentenolone[o]	Very strong, caramellic-maple, lovage	
Methyl-alpha-ionone	Orris, violet, woody, floral, oily with woody raspberry notes	*Dessert*
Methyl-thio-methyl-pyrazine	Roasted nut, burnt, meaty	*Dessert*
Orange oil	Orange	*Fruit, citrus*
Peppermint oil	Peppermint	*Menthol/mint*
Piperonal	Sweet, floral-cherry (heliotrope); sweet cherry-vanilla taste	
Propenyl guaethol	Sweet, vanilla, creamy, phenolic, anisic flavor	*Dessert*
Propionic acid	Pungent, sour milk, cheese	*Fruit, berries*
Trans-2-hexenal	Green, fruity, fresh, apple with leafy and grassy notes	*Fruit, other*
Trans-anethole	Sweet, herbaceous, anise (artificial licorice)	*Spices*
Triethyl citrate	Weak, sweet, winey-fruity-plum-like odor; slight bitter taste	*Nuts*
Vanillin[o]	Sweet, powerful, creamy, vanilla-like	
β-damascone	Blackcurrant, plum, rose, honey, tobacco	*Tobacco*
γ-Decalactone[o]	Coconut-peach	
γ-Dodecalactone	Fatty, fruity, peach odor	*Nuts*
γ-Hexalactone	Coumarin-like, sweet, creamy note	*Dessert*
γ-Undecalactone[o]	Strong fatty, peach-apricot	

* A flavoring ingredient added to more than 10% of the e-liquids in a single category was considered "category-specific".

º The 25 flavoring ingredients that were most frequently added to e-liquids from the EU-CEG dataset. These data are also presented in Table 1 of the main manuscript.

Supplementary Table S4

Table S4: Prediction accuracy of a random forest (RF) algorithm that assigns e-liquids a flavor category [1] based on information about flavoring composition. Actual category is the flavor category where we classified the product in based on flavor-related information from the EU-CEG system.

			Predicted category																		
	Category	Description	Cat01	Cat02	Cat03	Cat04	Cat05	Cat06	Cat07	Cat08	Cat09	Cat10	Cat11	Cat12	Cat13	Cat14	Cat15	Cat16	Total	Predicted correctly	Predicted at 2nd place	Remaining
Actual category	Cat01	tobacco	2323	22	0	0	0	0	1	2	0	1	8	11	0	18	0	0	2386	2323 (97%)	41 (2%)	22 (1%)
	Cat02	menthol/mint	336	726	0	0	0	0	1	3	3	0	2	4	0	1	0	0	1076	726 (67%)	324 (30%)	26 (2%)
	Cat03	nuts	64	3	54	0	0	0	0	0	0	2	4	13	0	20	0	0	160	54 (34%)	58 (36%)	48 (30%)
	Cat04	spices	75	6	0	67	0	0	1	3	0	0	1	9	0	0	0	0	162	67 (41%)	79 (49%)	16 (10%)
	Cat05	coffee/tea	213	0	0	0	163	0	0	0	4	4	4	13	0	5	0	0	406	163 (40%)	208 (51%)	35 (9%)
	Cat06	alcohol	112	3	0	0	2	197	1	2	5	1	7	10	0	6	0	0	346	197 (57%)	124 (36%)	25 (7%)
	Cat07	other beverages	189	1	0	0	1	4	512	15	9	5	8	26	0	3	0	0	773	512 (66%)	219 (28%)	42 (5%)
	Cat08	fruit (berries)	431	13	0	0	0	0	4	1392	1	5	13	25	1	11	0	0	1896	1392 (73%)	462 (24%)	42 (2%)
	Cat09	fruit (citrus)	189	9	0	0	0	0	9	24	413	6	1	6	0	2	0	0	659	413 (63%)	213 (32%)	33 (5%)
	Cat10	fruit (tropical)	293	5	0	0	0	0	2	13	5	639	15	14	1	6	0	0	993	639 (64%)	307 (31%)	47 (5%)
	Cat11	fruit (other)	586	20	0	0	1	3	4	22	15	7	1476	25	1	3	0	0	2163	1476 (68%)	610 (28%)	77 (4%)
	Cat12	dessert	270	2	0	0	6	2	1	6	11	2	17	1140	0	28	0	0	1485	1140 (77%)	314 (21%)	31 (2%)
	Cat13	candy	209	12	1	5	0	0	0	20	2	4	17	14	398	0	0	0	682	398 (58%)	256 (38%)	28 (4%)
	Cat14	other sweets	208	5	4	0	0	0	1	4	0	1	8	35	0	411	0	0	677	411 (61%)	232 (34%)	34 (5%)
	Cat15	other flavors	60	3	0	0	0	0	0	1	5	0	3	8	0	6	65	0	151	65 (43%)	61 (40%)	25 (17%)
	Cat16	unflavored	232	0	0	0	0	0	0	0	0	0	0	0	0	0	0	6	238	6 (3%)	232 (97%)	0 (0%)
	Cat17	unclassifiable	1447	189	3	3	14	67	27	232	51	41	257	173	4	74	4	0	2586	NA	NA	NA
		Total	7237	1019	62	75	187	273	564	1739	524	718	1841	1526	405	594	69	6	16839	9982 (70%)	3740 (26%)	531 (4%)

References

[1] Krusemann EJZ, Boesveldt S, de Graaf K*, et al.* An E-liquid Flavor Wheel: A Shared Vocabulary based on Systematically Reviewing E-liquid Flavor Classifications in Literature. *Nicotine Tob Res* 2018.

[2] Leffingwell & Associates. Flavor-Base 9 - Tobacco Version for Windows XP/Vista/7&8. 2013.