PAPER

Check for updates

Quantifying PG : VG ratio and nicotine content in commercially available e-liquids using handheld Raman spectroscopy†

Paul I. C. Richardson, [ID] Adam Burke, Nigel Gotts and Royston Goodacre [ID] *

Electronic cigarettes are a popular nicotine consumption product that have risen in popularity as an alternative to cigarettes. However, their recent meteoric rise in market size and various controversies have resulted in the analyses of e-liquid ingredients to be focused on powerful laboratory-based slow methods such as chromatography and mass spectrometry. Here we present a complementary technology based on Raman spectroscopy combined with chemometrics as a fast, inexpensive, and highly portable screening tool to detect and quantify the propylene glycol : glycerol (PG : VG) ratio and nicotine content of e-cigarette liquids. Through this, the PG : VG ratio of 20 out of 23 commercial samples was quantified to within 3% of their stated value, while nicotine was successfully quantified to within 1 mg g^{-1} for 16 out of 23 samples without the need for accurate knowledge of flavonoid composition. High linearity was also achieved when flavours were kept constant. Finally, the limitations of Raman spectroscopy are discussed, and potential solutions are suggested.

Introduction

E-cigarettes are a recently popularised nicotine ingestion product designed as an alternative to smoking tobacco and a smoking cessation aid.[1] E-cigarettes liquids (e-liquids) are mixtures of propylene glycol and vegetable glycerine in which nicotine and various flavour chemicals are dissolved.[2,3] These liquids are absorbed through a wick and heated to be inhaled as a vapour. In contrast to traditional cigarette smoking, the lack of combustion is said to decrease the quantity of various toxicants.[4] This, alongside their use as an aid in tobacco cessation,[1] has been a major aspect of the touted health benefits of e-cigarettes as compared to traditional cigarettes. The second major selling point of these products is the degree of customisation available: everything from the ratio of propylene glycol to glycerine and nicotine content to the flavour and intensity can be altered to suit the taste of the user. These factors have led e-cigarettes to quickly become extremely popular and lucrative products and this popularity is still growing rapidly, with a global market size of $22.45 billion and forecasted to reach $182.84 billion in 2030.[5]

The speed at which e-cigarettes rose to cultural prominence, however, did not allow for the touted health benefit

claims and risks to be fully substantiated and caused nuanced regulations to significantly lag behind, which in turn led to several controversies. One of these was a concern that these products were leading to youth access to nicotine or acting as a gateway to tobacco.[6] Several companies, most notably JUUL, were suspected of targeting teenagers[7–10] by marketing e-cigarettes as "cool",[11] through social media,[8,12] and offering flavours that might appeal to them,[13,14] while using their official position as cigarette replacement aids to circumvent standard regulations generally present for nicotine containing products. This led to various responses around the world[15] ranging from complete bans[16] to banning certain flavours,[17] the presence of vitamins or health-related additives[18] or nicotine concentrations above 20 mg g^{-1}, alongside voluntary self-regulation[19,20] from the industry. Additionally, concerns have been raised regarding the possibility of a variety of new dangerous chemicals entering the lungs as a result, ranging from various degradants; *e.g.*, poly-ethylene glycol or formaldehyde[21–23] to flavour chemicals that, while deemed safe when ingested in food products, cause concern when entering the lungs.[24,25] Despite the claim that vaping is "95% less harmful than traditional cigarettes"[26] being often repeated as fact, the original claim lacked "hard evidence" and evidence of the contrary is growing.[27] Finally, the degree to which e-cigarettes have a beneficial effect on cigarette use cessation[3,28] has been put in question, especially with the averred increased risk from dual use. In 2016, the FDA placed tobacco products under its jurisdiction and requested a variety of data regarding the chemical composition and stability to create a robust

Centre for Metabolomics Research, Department of Biochemistry, Cell and Systems Biology, Institute of Systems, Molecular and Integrative Biology, University of Liverpool, BioSciences Building, Crown St, Liverpool, L69 7ZB, UK. E-mail: roy.goodacre@liverpool.ac.uk

regulatory framework, while the EU and the UK created their own regulatory framework.

As regulations have caught up and quantitative data regarding compliance has been sent to various governmental agencies, chromatographic separation methods and mass spectrometry have been the analytical methods of choice for many producers.[29–31] This is perhaps not surprising as these are established and highly sensitive methods, able to quantify analytes down to ng g^{-1} levels.[15] However, when sample preparation and instrument requirements are considered, they are time-consuming, laboratory based and resource intensive, as well as requiring specialist knowledge to operate. As such, these are generally not suited for routine screening analysis. Other methods such as SERS[32] and NMR spectroscopy[33,34] have been used to quantify nicotine and dangerous by-products in e-liquids, however once again specialist knowledge, significant sample preparation and high up-front costs are often required.

Here, by contrast, we propose the use of portable Raman spectroscopy instruments to be used as a quick screening tool for quality control purposes in both regulatory and industrial settings. This method trades sensitivity for speed and ease of use, requiring virtually no sample preparation or specialist knowledge to operate, and obtaining measurements in <60 s. Furthermore, we investigate the use of these instruments to collect measurement through containers, thus allowing users to test commercial e-liquid products without the need to open them, potentially at the site of purchase.

Methods

Materials and sample preparation

Commercial samples (n = 27) were purchased from two online retailers. A range of nicotine concentrations were obtained for three representative flavour profiles: tobacco, menthol, and various fruit flavours. Extra empty pods were graciously provided by acquaintances who consume the product; these were thoroughly rinsed in copious amounts of double distilled water and dried prior to use. Hexane and dichloromethane were obtained from Acros chemicals (Acros Organics, Thermo Scientific, Geel, Belgium), HPLC-grade methanol was obtained from Fisher chemicals (Fisher Scientific UK Ltd, Loughborough, England), decane standard was obtained from Thermo Scientific, while dodecane, pentadecane, nonadecane, and docosane standards were obtained from Alfa Aesar (Alfa Aesar, Heysham, England). All other chemicals were obtained from Sigma Aldrich (Sigma Aldrich, Gillingham, England). Aliquots of propylene glycol (PG) and glycerol (VG) were mixed in ratios ranging from 95% PG to 95% VG in 5% increments, and homogenised using a Sokany EW-071 milk frother (Sokany, Yiwu City, Zhejiang Province, China) with a hook attachment for ~10 s. A series of nicotine solutions ranging from 1 to 20 mg g^{-1} were made in both PG and VG, and also homogenised in the same way. Aqueous solutions of nicotine at 5, 10, 15, and 20 mg mL^{-1} were made serially from a 35 mg mL^{-1} base/stock solution.

Raman spectroscopy

For vial measurements, a 2 mL aliquot of each sample was transferred to a 2 mL glass vial. The Raman spectra were obtained using a SnRI CBeX (Snowy Range, Laramie, Wyoming, USA) portable Raman spectrophotometer in vial mode, connected to and controlled *via* the dedicated software package. A laser at wavelength of 785 nm was used for illumination at a power of 50 ± 3 mW at the sample. The spectral range was 400–2300 cm^{-1} with one data-point per wavenumber, leading to a total of 1901 'bins' per spectrum.

For the PG : VG ratio dataset, five machine replicates were obtained per sample using a randomised collection order. A single 1 s acquisition was acquired for each technical replicate, leading to a total of 105 spectra. For the nicotine content datasets and commercial samples, concentration ranges for propylene glycol and glycerol were randomised and measured separately using five sets of five machine replicates. Five ×2.5 s acquisitions were gathered consecutively per sample and the randomised sequence of samples was repeated five times, leading to 25 spectra acquired per sample. This led to a total of 525 spectra per concentration curve and 675 spectra in total for the commercial samples.

Data processing and analysis

All raw spectra were transferred to Matlab R2020A (The MathWorks, Natick, USA), where the 1801–2300 cm^{-1} spectral range was removed, as there were no discernible peaks in this region. Then, a least squares baseline correction algorithm based on that of Eilers[35] (smoothing parameter: 1×10^4, asymmetry parameter: 0.01) was applied, and the 1300–1450 cm^{-1} was removed due to random signals (presumably fluorescence) from the glass vials. Finally, a standard normal variate (SNV) normalisation function was applied. Where relevant, principal component analysis (PCA)[36] was used to reduce the dimensionality of the data by grouping variables into uncorrelated principal components (PCs) which could be used to elucidate relationships between spectra. Quantification was performed using partial least squares regression (PLSR).[36] Training sets containing evenly distributed concentrations (0%, 10%, 20%, 30% through to 100% PG for the PG : VG quantification; and 0, 2, 4, 6 through to 20 mg g^{-1} nicotine for nicotine quantification) of each dataset and these were used to create multiple models using an increasing number of latent variables (LVs). Each model was tested using k-fold cross-validation (k = 100), and the model showing the lowest root-mean-squared error (RMS) of cross validation was used for quantification. Test sets comprised of the rest of the prepared samples (*i.e.* 5, 15 … 95%PG; or 1, 3 … 19 mg g^{-1} nicotine), along with the commercial samples.

Results and discussion

PG : VG ratio

E-cigarette liquid is generally composed of four basic components: a base consisting of a propylene glycol/glycerol (from

vegetable source, hence VG) mixture, water, nicotine, and flavouring additives.[2,3] Predictably, as the PG : VG base is the largest component it has the highest impact on the Raman spectrum; as such, it must be considered first. Not all e-liquids use the same ratio of PG : VG as a base; due to the advertised customisability and advent of "artisanal" e-liquids manufacturers, similar to microbreweries in the craft beer market, e-liquids can be found with ratios ranging from 100% PG to 100% VG.

Glycerol is more viscous and has a sweeter taste, perceived to be approximately half that of sucrose (table sugar).[37] Higher proportions of it in e-liquids are associated with a more mellow flavour, a less intense throat-feel and larger exhaled clouds. These qualities can be attractive to people who are interested in the aesthetic aspect of vaping.[38] However, the higher viscosity can lead to obstructions within the vaping equipment and thus will often require more maintenance. Propylene glycol, on the other hand, is an odourless and less viscous option.[39] It is a better carrier of flavour than glycerol; as such, higher proportions tend to give the e-liquid a more intense flavour.[38] Additionally, higher proportions of it are associated with a harder "hit" at the back of the throat,[40] a sensation that can be appealing to tobacco smokers using a transition to vaping as a means to stop smoking traditional cigarettes. PG is also better absorbed through the wick and less likely to congest the tube. However, some find it to be an irritant.[38] These differences in properties and sensibilities allow for various base concentrations to thrive in the market.

Fig. 1 depicts the changes in Raman spectra as the base mixture changes from 100% glycerol (in blue) to 100% propylene glycol (in red). Unsurprisingly, despite the similar size and functional groups, these different molecules lead to fundamentally different spectra, meaning various ratios can be readily differentiated from each other at a glance.

While Fig. 1A shows the whole examined spectral range, Fig. 1B is a zoom on the 460–550 cm^{-1} region. Two unique peaks are presented: one at 485 cm^{-1} belonging to glycerol and one at 522 cm^{-1} belonging to propylene glycol, each resulting from their respective molecule's δ(CCO) band.[41] These two peaks, while close, are clearly defined, with little baseline overlap and unique, thus presenting an opportunity to simplify the task of quantification.

Although the use of more points of difference generally leads to superior differentiation, it is important to remember that the samples used are incomplete models containing neither nicotine or flavourings. The addition of these chemicals could potentially yield extra peaks, which in turn could limit the model's robustness. As it is not feasible to test every single combination of flavour and nicotine concentration, this potential issue was mitigated by minimising the spectral range tested, thus decreasing the probability of so called 'irrelevant' peaks affecting the model. Fig. 1C and D plot example results of PLSR modelling using the full 400–1800 and 462–550 cm^{-1} spectral range, respectively. Even using a shorter range, the model is clearly able to predict concentrations (red circles) of samples in a test set with excellent accuracy. Fig. 1E demon-
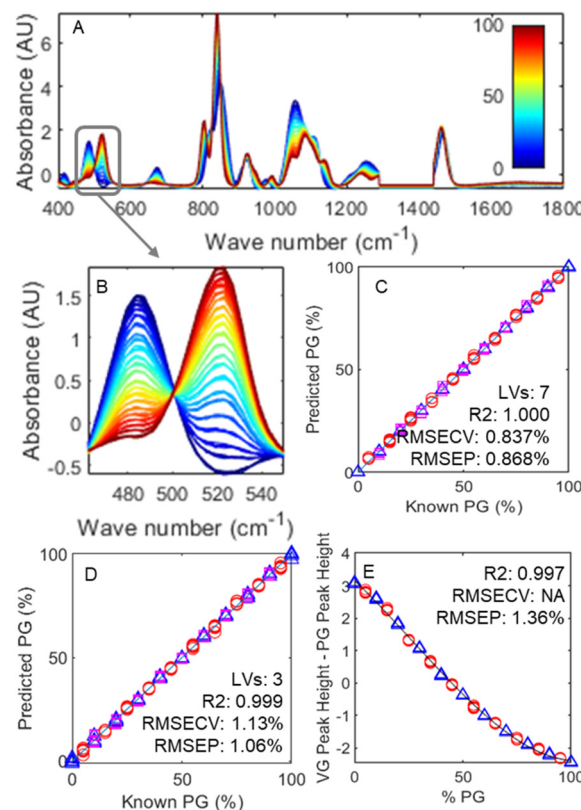
Fig. 1 A: Baseline corrected and normalised Raman spectra of PG : VG mixtures ranging from 0 : 100 PG : VG (blue) to 100 : 0 PG : VG (red). The grey box in A represents the portion depicted in B, which shows the two peaks used in predictive models. C and D: Plotted results for the PLSR model made using the 400–1800 cm^{-1} and 462–550 cm^{-1} spectral ranges respectively: blue triangles, pink squares and red circles represent the training, cross-validation and test sets respectively. E: Plotted difference between the normalised peak heights at 485 and 522 cm^{-1}, alongside a cubic regression ($y = 4.234 \times 10^{-6}x^3 - 3.227 \times 10^{-4}x^2 - 0.06451x + 3.165$): blue triangles and red circles represent training and test sets respectively. LVs represent the number of latent variables that gave the optimal model; R^2 represents the coefficient of linearity for the training sets; while RMSECV and RMSEP depict the root mean square error of prediction for the cross-validation and test sets in terms of %PG.

strates the results of taking this simplification further by only using the difference between the scaled peak intensities at 485 cm^{-1} and 522 cm^{-1}. Due to the flattening out of the curve at the extremes of the PG : VG ratio, a cubic function was chosen as a best fit model.

Unsurprisingly, the model with the highest predictive power was the one utilising the whole 400–1800 cm^{-1} spectral range, which achieved an RMSEP of 0.87%, followed by the 462–550 cm^{-1} range which achieved an RMSEP of 1.06%, and finally the peak heights model achieving an RMSEP of 1.36%. To put these numbers in perspective, commercial e-liquids are generally found in ratio increments of 5–10%; as such, even the highest RMSEP obtained using the peak heights model is sufficient to estimate accurately the PG : VG ratio. The ability to use this simpler model is beneficial not only as it significantly reduces complexity and thus the model is more parsi-

monious[42] as well as a reduction in the computational power needed, but also minimises the risk of flavour peaks, which have thus far been ignored, interfering with the prediction ability.

Nicotine concentration

After propylene glycol and glycerol, nicotine is the most common component in e-liquids. Indeed, while many nicotine-free products exist, the product was originally conceived as less harmful alternative to cigarettes.[43] As such, the vast majority of products will contain nicotine, generally in its free-base form. Fig. 2 shows the changes in Raman spectra resulting from the addition of nicotine in both PG and VG.

One of the major difficulties of creating a robust model for the quantification of nicotine in e-liquids can be seen by comparing its addition in VG (Fig. 2A) and PG (Fig. 2D). The visible change in spectra largely originates from two areas: the area at 1030 cm^{-1} attributed to a combination of nicotine's two most prominent peaks, and an area at 1600 cm^{-1} which, while far weaker, is situated in an area where both PG and VG are silent. All of these peaks originate from the pyridine moiety vibration bands ν(CC) and ν(CN) stretching, and ρ(CH) rocking.[44]

The scale of change in spectra resulting from the addition of nicotine (black to blue and red respectively) compared to the spectral differences originating from the PG : VG baseline is extremely small, as exemplified by the zoomed in sections overlaid in the figures. However, the similarity between the two PC 1 loadings plots (Fig. 2B & E) shows that while the baseline effect is very different, the differences between spectra remain largely consistent. Fig. 2G, comparing the normalised and averaged Raman spectra for nicotine in water (black) and the 20 mg mL^{-1} nicotine in PG and VG sample after subtraction of their respective solvent only spectra (red and blue respectively) with their respective PC 1 loadings plots further confirm that the main factor allowing the separation of these spectra is the change in nicotine signal. There is therefore potential for these baseline discrepancies to be removed and for a robust universal predictive model to be created.

Table 1 shows the nicotine prediction results obtained using PLS regression models. To create these models, spectra were first separated into a training set containing all samples with even nicotine concentration and a test set containing all samples with odd nicotine concentrations (as described in the Materials section). In addition, a k-fold ($k = 100$) cross validation method was used to find the best number of latent variables (LVs) to use, chosen with the aim of minimising the RMSECV. Alongside the PG and VG sets, two combined sets were created using the training sets (labelled as trn) from either PG or VG and the test set (tst) from the other. Additionally, the 0 and 20 mg g^{-1} spectra of the test set solvent ratio were added to the training set; these solvent ratio "anchors" allowed the model to account for the significant spectral differences originating from the solvent and drastically improved the model prediction ability.

The RMSEP values for each model show the ability for PLSR to quantify the concentration of nicotine in samples in both ideal and imperfect situations (*i.e.*, those where there is a difference in the base e-liquid). Even the least accurate model out of the four, which used a PG training set and a VG test set,
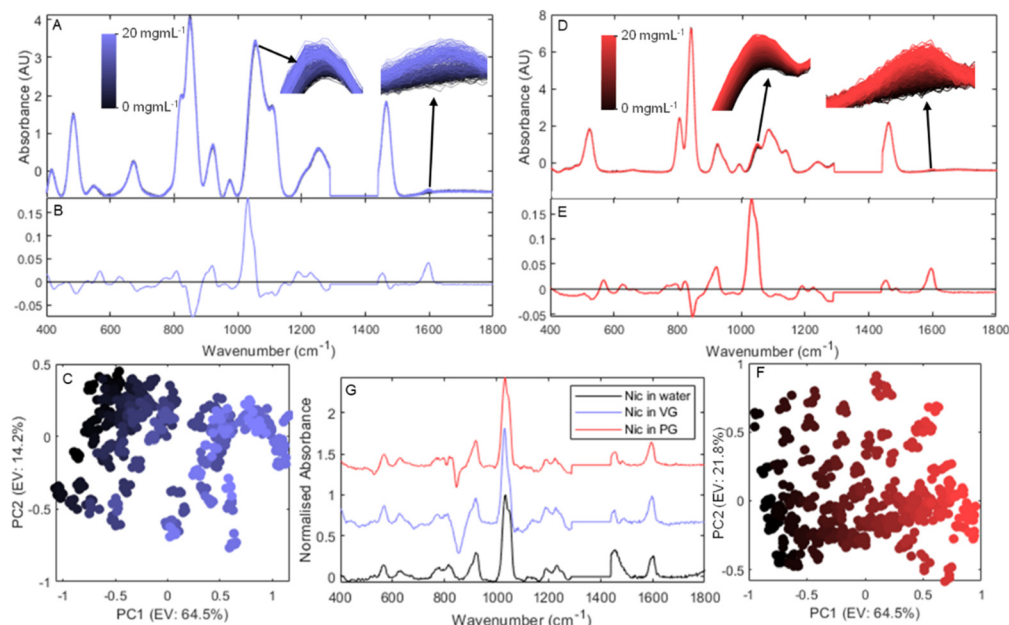
Fig. 2 Results of Raman spectroscopic analysis of nicotine in VG (blue) and PG (red). A and D depict the baseline corrected and normalised spectra of 0–20 mg g^{-1} nicotine in VG (black to blue) and PG (black to red) respectively. C and F depict PCA scores plots for these spectra (EV in parenthesis = explained variance), while B and E depict the loadings for PC1 for each dataset and highlight important regions related to increasing levels of nicotine. G compares these PC loadings plot to a baseline corrected and normalised spectrum of a 20 mg mL^{-1} aqueous nicotine solution.

Table 1 Tabulated results for the PLSR and regression models for the quantification of nicotine content in homemade e-liquid samples. LVs represents the number of latent variables that gave the optimal model where trn and tst refer to samples in the training or test sets, respectively; R^2, Q^2CV and Q^2P represent the coefficient of linearity for the training, cross-validation and test sets, respectively; while RMSEC, RMSECV and RMSEP depict the root mean square error of prediction for each set in terms of mg g^{-1} of nicotine

Model	LVs	R^2	Q^2CV	Q^2P	RMSEC (mg g^{-1})	RMSECV (mg g^{-1})	RMSEP
PG	4	0.999	0.997	0.998	0.236	0.307	0.268
VG	4	0.991	0.981	0.991	0.573	0.905	0.524
PG-trn, VG-tst	6	0.998	0.994	0.965	0.303	0.389	1.08
VG-trn, PG-tst	7	0.995	0.979	0.991	0.490	0.756	0.546

only reaches an error prediction of 1 mg g^{-1}. This would be capable of detecting even low concentrations of nicotine in commercial products, which rarely fall below 3 mg g^{-1}, but also would be able to clearly distinguish between different concentrations in products, which tend to change in 3 mg g^{-1} or larger increments. It is also worth mentioning that this apparent weakness is likely due to imperfect mixing rather than a weakness in the model. In particular the high viscosity of glycerine made ensuring a completely homogenous nicotine solution extremely difficult, leading to the use of a portable milk frother which, while greatly improving the mixing, may not have been entirely perfect. The most accurate model, which used only the PG data set, showed excellent accuracy with an RMSEP of 0.2684 mg g^{-1}, which would be more than sufficient to consistently detect 1 mg g^{-1} nicotine in products, the lowest concentration found on sale. Overall, these results show that, even in suboptimal conditions, Raman spectroscopy has potential to be an effective screening method for nicotine detection.

Commercial samples

Alongside varying ratios of PG and VG in commercial e-liquids, a significant analytical challenge is the presence of flavour additives that can affect the spectra in a variety of ways. Indeed, a multitude of compounds have been found in e-liquids at varying concentrations.[45] To add to the challenge, each company will use their own proprietary recipes for each flavour, for which the specific compounds and concentrations are not publicly described. As such, even with general *a priori* knowledge based on the Raman signatures of the most common compounds and likely main components of each flavour, significant work would be necessary to create what would be, at best, a vague facsimile of the product to be tested.

A total of 27 different products from two different producers were purchased and tested. The products were chosen to contain both varying concentrations of nicotine and PG : VG ratio, and three flavour categories were tested: menthol, tobacco, and fruity. This spread aimed to study a varied range of popular products, while allowing for the possibility of understanding why certain, if any, products could not be quantified accurately. Additionally, we aimed to assess the possibility of quantifying these products through-container (in the commercial bottle) using the spectrometer's point-and-shoot mode alongside the use of vials. As such, the two companies were chosen for their different containers, with one supplying

refill bottles while the other supplying disposable pods that slot into compatible e-cigarettes.

While acquiring data, one concern that became apparent was the presence of a higher baseline present in many commercial samples, likely occurring as a result of the samples' colouring. This could often be removed during data pre-processing; however, some backgrounds were high enough to saturate the detector at the stated measurement times, and are likely to be due to fluorescence. Attempts to reduce the acquisition time per spectrum and manually add several spectra together to normalise acquisition times led to significantly lower signal-to-noise ratios due to the weakness of the nicotine signal relative to others. Therefore, these were omitted from analysis. Also omitted were the in-container measurements for the pods sold by producer B. Due to the size and configuration of the pod, which is angled, very thin and contains a metal rod in the centre, we did not succeed in obtaining spectra of sufficient quality or consistency for analysis.

Fig. 3C depicts the average results for the analysis of commercial samples when analysed in vials. For the majority of the products, the predicted percentage of PG in the solvent matches the stated values, often deviating by no more than 3%. These were obtained using the difference in scaled peak heights method outlined earlier. Interestingly, while the other methods obtained better quantification in optimal conditions, the presence of new peaks and background information in the commercial samples led to more incorrect results reinforcing that a more parsimonious model is preferred. There remain, however, three outliers within producer B's samples which show significant deviation from the stated values; the fruit 1, menthol 2 and tobacco samples show significant underestimations of 15.2%, 23.7% and 17.5% respectively. However, further inspection of these spectra revealed them to match their predicted PG : VG ratio far better than their stated one (Fig. S1†). Producer B does not state the ratio in their products anywhere in the packaging or on their website – this information was obtained from their customer service team. As such, both miscommunication or an error on their part are possible reasons for this discrepancy.

Although the ability to quantify the PG : VG ratio is an important pre-requisite for further probing of e-liquids using Raman spectroscopy, the ability to screen for nicotine concentration in products is likely to be more attractive to both producers, regulators and users – especially as altering PG : VG can lead to a false 'hit' that may be associated by the user with
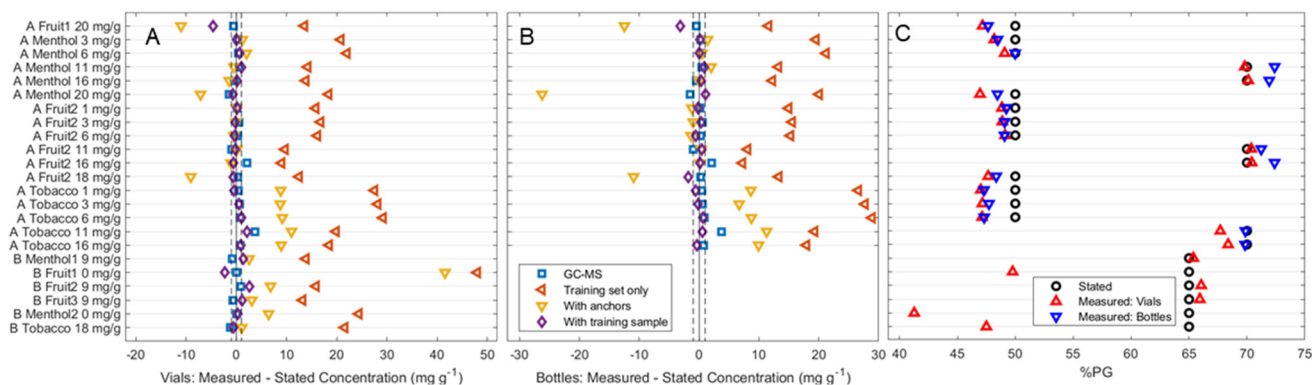
Fig. 3 Predicted nicotine concentrations of each tested commercial product relative to their stated concentration, using different methods and different training sets. Blue squares refer to results obtained using GC-MS. All others were obtained using Raman spectroscopy followed by PLSR: orange left-facing triangles use only the nicotine in PG training set, yellow downward-facing triangles add to this anchors at the stated PG : VG ratio, and purple diamonds add both anchors and one commercial sample to the training set. Dotted line represents ±1 mg g^{-1} from the commercially stated concentration. C compare the PG : VG ratio (as %PG) stated by manufacturers (black circles) to the calculated values using vial and bottle measurements (red upward-facing and blue downward-facing triangles respectively). A tabulated version of the "with training sample" results (purple diamonds, 3A/B) and the %PG results is available as Table 2 and a complete tabulation of 3A/B is present in the ESI (Table S1†), alongside the experimental details for the GC-MS data.

higher nicotine levels. Fig. 3A and B depict the predicted concentrations for each commercial sample relative to their stated concentration when using different methods. These consist of a ground truth obtained using GC-MS – blue squares – alongside three different statistical models performed on the Raman spectra: red left-facing triangles where only the previous nicotine in PG training set was used; the yellow downward-facing triangles depicting the addition of 0 and 20 mg g^{-1} anchors at the stated PG : VG ratio; and purple diamonds where alongside the anchors, one other commercial sample was added to the training set. A tabulated form of this figure which is expanded in terms of details includes the specific commercial sample used in each model's training set can be found in the ESI.†

One can observe that increasing the amount of information leads to a general increase in accuracy of the Raman predictions. All predicted concentrations obtained using only the nicotine in PG training set drastically overestimate the concentration of nicotine. The addition of anchors at the correct PG : VG ratio improve the accuracy, and in the case of most of producer A's menthol and second fruit-flavoured products find an accurate value. However, this accuracy is inconsistent – many products still overestimate the nicotine concentrations, while producer A's 18+ mg g^{-1} samples all underestimate their value. While the failure to correctly quantify the nicotine concentrations of tobacco samples is likely due to the presence of flavour chemicals overlapping with the 1600 cm^{-1} nicotine peak (Fig. S3†), the underestimation of the highest concentration samples may also be affected by the presence of protonated nicotine salts in addition to free-base nicotine. There are two basic nitrogen groups in nicotine (pK_{a_1} = 3.12, pK_{a_2} = 8.02) allowing for three possible forms: free-base, monoprotonated and diprotonated. While the diprotonated form is not considered safely obtainable in e-liquids, the ratio of free-base

and monoprotonated is often manipulated through the addition of acidic solutions (*e.g.* benzoic acid) or ammonia.[33,46] The pH of e-liquids is thus a relevant manufacturing concern especially in e-liquids with higher nicotine content,[47] as protonated salts both affect the bioavailability of nicotine and lead to a less harsh sensation,[48] which can allow for tolerance of higher concentrations.

pH also has a small effect on the nicotine spectrum – when measured in a solvent with a minimal Raman contribution (*e.g.* water), the 1030 cm^{-1} signal discussed previously is actually combined with a second peak at 1050 cm^{-1} (Fig. S4A†). These are inversely correlated to each other with respect to pH. Due to interference from the PG : VG matrix in e-liquids, the nicotine signal only presents itself as an increase in the 1045 cm^{-1} peak area rather than having a peak of its own (Fig. S4B†).

Excluding the use of nicotine salts, the two primary factors affecting the pH of e-liquids are the flavour profile, which provides an e-liquids pH baseline, and the nicotine concentration, where increases correlate with increases in pH. While the former does not affect the algorithm's predictive ability – producer A's fruit 2 and menthol flavour series have pH values of 5–7 and 8–9 respectively (Table S1†). Despite these complications both are accurately modelled – the latter will shift the influence of the nicotine peak towards 1030 cm^{-1} in a manner similar to our training set. However, the introduction of nicotine salts necessitates a decrease in pH, which in turn will move the nicotine signal's influence back towards 1050 cm^{-1} and lead to an underestimation of the nicotine concentration by the algorithm. The quantification of nicotine in salt form would therefore likely need further development and perhaps its own training set.

It is only by adding commercial samples, and thus spectral information related to the flavour compounds present, that

models are able to consistently obtain accurate results. There are, nonetheless, caveats that must be stated when discussing the use of commercial samples in the training set. To find the results in Fig. 3 and Table 2, the PLSR algorithm was run using each commercial sample in the training set one at a time, and the optimal model was chosen based on its closeness to the stated nicotine content. Furthermore, while the commercial sample closest in profile to the one being tested was generally given priority when several models were close in predictive ability (*e.g.* same PG : VG ratio, brand and flavour), there were several instances where the most accurate model was unexpectedly different. This is obviously not a method that could be replicated in a regulatory setting; further development into a single coherent model would be necessary, perhaps using multiple samples in the training set. It should be noted, however, that additives do not negatively affect results when accounted for; spiking nicotine into commercial e-liquid led to a linear and predictable increase in the 1045 cm^{-1} peak height (Fig. S5†). Additionally, it is clear that results acquired through producer A's bottle are once more closely comparable to those obtained in vials, further showing promise for the use of portable Raman instruments as a useful tool in the field.

To recapitulate, while there is strong potential for Raman spectroscopy to work as an on-site screening tool, three limitations must first be overcome. First, components present in some e-liquids lead to a high fluorescence background, overloading the signal; work must be performed to mitigate this issue, and could be based on redesigning the instrument to use time-gated measurements or collect anti-Stokes signals. Second, a consistent model must be set up to account for the varied flavour profiles in e-liquids without relying on matching

it to the stated nicotine content, and for deployment a large database with baseline measurements based on (*e.g.*) GC-MS would likely achieve this. As such a model would better include flavours in its analysis, it also has potential to increase the model's accuracy. Third, while bottles lend themselves to the point-and-shoot system well, the shape and metal rod in the pods made them extremely difficult to analyse – Raman spectrometer hardware would likely need bespoke adaptation or an add-on for through-container analysis of these pods. Finally, it is important to note that while Raman spectroscopy has significant speed, cost, and ease of use advantages when compared to the gold standard chromatographic methods, we view its potential as complementary rather than supplanting. An ideal framework would involve Raman analysis as a routine first point of call in the field by regulatory agents or as a quality confirmation tool by manufacturers, comparing new batches with expected spectra from previously produced batches. Any samples arousing suspicion being sent for further analysis using chromatographic methods, thereby increasing efficiency in the quality control process.

Conclusion

Here, we present Raman spectroscopy as a fast, portable and simple screening tool for in-container measurements of the PG : VG ratio and nicotine concentration of e-liquids. Using a simplified model without flavours, Raman spectroscopy was able to predict the PG : VG ratio of test samples correctly typically with a root means squared error of within 0.9% using a PLSR model using entire spectra (from 400–1800 cm^{-1}), and within 1.4% using just the difference between the peak heights at 485 and 522 cm^{-1}. Furthermore, using PLSR modelling, we were able to predict the nicotine concentration in 100% PG using a 100% PG training set to within 0.27 mg g^{-1} RMS error, and using a 100% VG training set with 100% PG anchors to 0.55 mg g^{-1} RMS error. Twenty-seven commercial samples from two different producers were also tested spanning multiple flavours and nicotine concentrations. For these, PG : VG was consistently predicted within 3% of the stated value using the difference between the two unique peaks, bar three outliers from producer B which may be explained by a miscommunication when obtaining the concentration from the producer's customer service team, as the PG : VG ratio was not stated on the product. Additionally, while the use of the previous training set and anchors at the correct PG : VG ratio was often sufficient, the addition of one commercial sample to the training set of the PLSR model led to significantly higher consistency, with 12 out 23 samples predicting a concentration within 1 mg g^{-1} of the stated concentration and 19 out of 23 samples reaching a concentration within 1.5 mg g^{-1} of the stated concentration. The use of point and shoot mode on the bottled sampled samples led to comparable results for both PG : VG and nicotine quantification. We also present technical limitations and currently feasible options to overcome them in the future.

Table 2 Tabulated best results Fig. 3. For nicotine, "vials" and "bottles" describe the predicted nicotine concentration in mg g^{-1} for the models including a sample in their training set using measurements performed in vial mode and point-and-shoot mode on containers

Producer A	Menthol	3	3.08	3.1	50	48.2	48.5
		6	6.67	6	50	49.1	50
		11	12.05	11.8	70	69.8	72.4
		16	16.11	16.3	70	70.1	71.9
		20	19.26	21.1	50	47	48.5
	Fruit 2	1	1.1	0.8	50	48.8	49.2
		3	2.81	3.3	50	48.9	49.1
		6	5.61	5.42	50	49.1	49.1
		11	10.83	11.4	70	70.4	71.2
		16	15.38	16.1	70	70.4	72.4
		18	17.41	16.2	50	47.7	48.4
	Tobacco	1	0.57	0.4	50	47	47.3
		3	3.46	2.8	50	47.2	47.7
		6	6.97	6.6	50	47.1	47.3
		11	13.15	11.5	70	67.7	69.9
		16	17.41	15.62	70	68.4	69.8
Producer B	Menthol 1	9	10.38		65	65.4	
	Fruit B1	0	−2.31		65	49.8	
	Fruit B2	9	11.65		65	66	
	Fruit B3	9	10.17		65	65.9	
	Menthol 2	0	0.32		65	41.3	
	Tobacco 1	18	17.35		65	47.5	

Author contributions

PICR contributed towards conceptualisation, investigation, methodology, validation and writing the original draft. AB and NG contributed towards methodology and validation. RG contributed towards Funding acquisition, supervision and reviewing and editing the publication.

Conflicts of interest

The authors declare no conflicts of interest.

Acknowledgements

References

1 D. Ramamurthi, P. A. Gall, N. Ayoub and R. K. Jackler, Leading-Brand advertisement of quitting smoking benefits for E-Cigarettes, *Am. J. Public Health*, 2016, **106**(11), 2057–2063, DOI: 10.2105/AJPH.2016.303437.

2 G. S. Helen and D. L. Eaton, *Public Health Consequences of E-Cigarette Use*, vol. 178. 2018. DOI: 10.1001/jamainternmed.2018.1600.

3 C. Franck, K. B. Filion, J. Kimmelman, R. Grad and M. J. Eisenberg, Ethical considerations of e-cigarette use for tobacco harm reduction, *Respir. Res.*, 2016, **17**(1), 1–9, DOI: 10.1186/s12931-016-0370-3.

4 M. L. Goniewicz, M. Gawron, D. M. Smith, M. Peng, P. Jacob and N. L. Benowitz, 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**(2), 160–167, DOI: 10.1093/ntr/ntw160.

5 Grand View Research. *E-Cigarette And Vape Market Size, Share & Trends Analysis Report By Product (Modular Devices, Rechargeable), By Distribution Channel (Online, Retail), By Region (APAC, North America), And Segment Forecasts, 2023–2030*; 2022. https://www.grandviewresearch.com/industry-analysis/e-cigarette-vaping-market.

6 O. A. Wackowski, D. P. Giovenco, B. Singh, M. J. Lewis, M. B. Steinberg and C. D. Delnevo, Content analysis of US news stories about e-cigarettes in 2015, *Nicotine Tob. Res.*, 2018, **20**(8), 1015–1019, DOI: 10.1093/ntr/ntx170.

7 D. M. Vallone, M. Bennett, H. Xiao, L. Pitzer and E. C. Hair, Prevalence and correlates of JUUL use among a national sample of youth and young adults, *Tob. Control*, 2019, **28**(6), 603–609, DOI: 10.1136/tobaccocontrol-2018-054693.

8 M. Al-Hamdani, D. B. Hopkins, A. Hardardottir and M. Davidson, Perceptions and Experiences of Vaping Among Youth and Young Adult E-Cigarette Users: Considering Age, Gender, and Tobacco Use, *J. Adolesc. Health*, 2021, **68**(4), 787–793, DOI: 10.1016/j.jadohealth.2020.08.004.

9 J. Liu, C. Vázquez-Otero, M. L. Berman and E. M. Stevens, Youth-appealing features in popular e-cigarette brand advertising in the USA after heightened scrutiny in 2018, *Tob. Control*, 2021, **32**(4), 497–500, DOI: 10.1136/tobaccocontrol-2021-056720.

10 O. Dyer, E-cigarette maker Juul will pay $ 462m to settle deceptive marketing allegations in six US states, *Br. Med. J.*, 2023, (381), 859, DOI: 10.1136/bmj.p859.

11 E. Brett, R. Krissinger and A. King, The rise and fall of e-cigarette cloud chasing appealing to youth, *Prev. Med. Rep.*, 2021, **24**, 101644, DOI: 10.1016/j.pmedr.2021.101644.

12 J. M. Alpert, H. Chen, H. Riddell, Y. J. Chung and Y. A. Mu, Vaping and Instagram: A Content Analysis of e-Cigarette Posts Using the Content Appealing to Youth (CAY) Index, *Subst. UseMisuse*, 2021, **56**(6), 879–887, DOI: 10.1080/10826084.2021.1899233.

13 D. R. Davis, M. E. Morean, K. W. Bold, *et al.*, Cooling e-cigarette flavors and the association with e-cigarette use among a sample of high school students, *PLoS One*, 2021, **16**(9 September), 1–8, DOI: 10.1371/journal.pone.0256844.

14 L. M. Schneller, M. Bansal-Travers, M. L. Goniewicz, S. McIntosh, D. Ossip and R. J. O'connor, Use of flavored E-cigarettes and the type of E-cigarette devices used among adults and youth in the US—results from wave 3 of the population assessment of tobacco and health study (2015–2016), *Int. J. Environ. Res. Public Health*, 2019, **16**, 2991–3002, DOI: 10.3390/ijerph16162991.

15 A. Taylor, K. Dunn and S. Turfus, A review of nicotine-containing electronic cigarettes—Trends in use, effects, contents, labelling accuracy and detection methods, *Drug Test. Anal.*, 2021, **13**(2), 242–260, DOI: 10.1002/dta.2998.

16 Global Tobacco Control. Country Laws Regulating E-Cigarettes. Published 2020. https://globaltobaccocontrol.org/en/policy-scan/e-cigarettes/sale.

17 U.S. Food and Drug Administration. FDA finalizes enforcement policy on unauthorized flavored cartridge-based e-cigarettes that appeal to children, including fruit and mint. Published 2020. Accessed January 20, 2023. https://www.fda.gov/news-events/press-announcements/fda-finalizes-enforcement-policy-unauthorized-flavored-cartridge-based-e-cigarettes-appeal-children.

18 MHRC. E-cigarettes: regulations for consumer products. Published 2016. Accessed January 20, 2023. https://www.gov.uk/guidance/e-cigarettes-regulations-for-consumer-products.

19 Juul Labs. JUUL LABS ACTION PLAN. JUUL Labs Newsroom. Published 2018. Accessed January 20, 2023. https://www.juullabs.com/juul-labs-action-plan/.

20 Philip Morris International. Marketing standards. Published 2021. Accessed January 20, 2023. https://www.pmi.com/our-views-and-standards/standards/marketing-standards.

21 R. M. Strongin, E-Cigarette Chemistry and Analytical Detection, *Annu. Rev. Anal. Chem.*, 2019, **12**(1), 23–39, DOI: 10.1146/annurev-anchem-061318-115329.E-Cigarette.

22 Q. Wang, X. Ji and I. Rahman, Dysregulated metabolites serve as novel biomarkers for metabolic diseases caused by e-cigarette vaping and cigarette smoking, *Metabolites*, 2021, **11**(6), 230–246, DOI: **10.3390/metabo11060345**.

23 N. R. Jaegers, W. Hu, T. J. Weber and J. Z. Hu, Low-temperature (< 200 °C) degradation of electronic nicotine delivery system liquids generates toxic aldehydes, *Sci. Rep.*, 2021, **11**(1), 1–12, DOI: **10.1038/s41598-021-87044-x**.

24 J. Gerloff, I. K. Sundar, R. Freter, *et al.*, Inflammatory Response and Barrier Dysfunction by Different e-Cigarette Flavoring Chemicals Identified by Gas Chromatography–Mass Spectrometry in e-Liquids and e-Vapors on Human Lung Epithelial Cells and Fibroblasts, *Appl. In Vitro Toxicol.*, 2017, **3**(1), 28–40, DOI: **10.1089/aivt.2016.0030**.

25 S. N. Langel, F. L. Kelly, D. M. Brass, *et al.*, E-cigarette and food flavoring diacetyl alters airway cell morphology, inflammatory and antiviral response, and susceptibility to SARS-CoV-2, *Cell Death Discovery*, 2022, **8**(1), 64–75, DOI: **10.1038/s41420-022-00855-3**.

26 A. McNeill, L. S. Brose, R. Calder, L. Bauld and D. Robson, *Evidence Review of E-Cigarettes and Heated Tobacco Products 2018. A Report Commissioned by Public Health England*, 2018. **https://www.gov.uk/government/publications/e-cigarettes-and-heated-tobacco-products-evidence-review**.

27 T. Eissenberg, A. Bhatnagar, S. Chapman, S. E. Jordt, A. Shihadeh and E. K. Soule, Invalidity of an oft-cited estimate of the relative harms of electronic cigarettes, *Am. J. Public Health*, 2020, **110**(2), 161–162, DOI: **10.2105/AJPH.2019.305424**.

28 R. J. Wang, S. Bhadriraju and S. A. Glantz, E-cigarette use and adult cigarette smoking cessation: A meta-analysis, *Am. J. Public Health*, 2021, **111**(2), 230–246, DOI: **10.2105/AJPH.2020.305999**.

29 Cooperation Centre for Scientific Research Relative to Tobacco. Recommended Method no. 84 Determination of Glycerin, Propylene Glycol, Water, and Nicotine in the Aerosol of E-Cigarettes by Gas Chromatographic Analysis. Published 2017. **https://www.coresta.org/sites/default/files/technical_documents/main/CRM_84-Oct2021.pdf**.

30 Cooperation Centre for Scientific Research Relative to Tobacco. Recommended Method no. 102 Determination of Tobacco- Specific Nitrosamines in E-Liquid by LC-MS/MS. Published 2023. **https://www.coresta.org/sites/default/files/technical_documents/main/CRM_102-Jan2023.pdf**.

31 Cooperation Centre for Scientific Research Relative to Tobacco. Cooperation Centre for Scientific Research Relative to Tobacco E-Vapour Sub-Group CORESTA Recommended Method no. 96 Determination of Formaldehyde and Acetaldehyde in E-Vapour Product AEROSOL. Published 2021. **https://www.coresta.org/sites/default/files/technical_documents/main/CRM_96-Feb2021.pdf**.

32 J. Y. Chien, Y. C. Gu, H. M. Tsai, *et al.*, Rapid identification of nicotine in electronic cigarette liquids based on surface-enhanced raman scattering, *J. Food Drug Anal.*, 2020, **28**(2), 302–308, DOI: **10.38212/2224-6614.1064**.

33 A. K. Duell, J. F. Pankow and D. H. Peyton, Free-Base Nicotine Determination in Electronic Cigarette Liquids by

1H NMR Spectroscopy, *Chem. Res. Toxicol.*, 2018, **31**(6), 431–434, DOI: **10.1021/acs.chemrestox.8b00097**.

34 P. J. Kerber, NMR Analyses of Flavored E-cigarette Liquids Before and After Aerosolization. Published online 2022.

35 P. H. C. Eilers and H. F. M. Boelens, Baseline Correction with Asymmetric Least Squares Smoothing, 2005 Unpublished Manuscript.

36 P. S. Gromski, H. Muhamadali, D. I. Ellis, *et al.*, A tutorial review: Metabolomics and partial least squares-discriminant analysis - a marriage of convenience or a shotgun wedding, *Anal. Chim. Acta*, 2015, **879**, 10–23, DOI: **10.1016/j.aca.2015.02.012**.

37 H. R. Moskowitz, Ratio scales of sugar sweetness, *Percept. Psychophys.*, 1970, **7**(5), 315–320, DOI: **10.3758/BF03210175**.

38 Vaping360 Team. PG *vs.* VG: What They Are and How to Use Them. Published 2022. Accessed January 26, 2023. **https://vaping360.com/learn/vg-vs-pg/?q=/pg-vs-vg-what-is-the-difference-and-what-should-i-use/&**.

39 K. Woelfel and T. G. Hartman, Mass Spectrometry of the Acetal Derivatives of Selected Generally Recognized as Safe Listed Aldehydes with Ethanol, 1,2-Propylene Glycol and Glycerol, *ACS Symp. Ser.*, 1998, **705**, 193–210, DOI: **10.1021/bk-1998-0705.ch017**.

40 T. T. Smith, B. W. Heckman, A. E. Wahlquist, K. M. Cummings and M. J. Carpenter, The Impact of E-liquid Propylene Glycol and Vegetable Glycerin Ratio on Ratings of Subjective Effects, Reinforcement Value, and Use in Current Smokers, *Nicotine Tob. Res.*, 2020, **22**(5), 791–797, DOI: **10.1093/ntr/ntz130**.

41 M. De Veij, P. Vandenabeele, T. De Beer, J. P. Remon and L. Moens, Reference database of Raman spectra of pharmaceutical excipients, *J. Raman Spectrosc.*, 2009, **40**(3), 297–307, DOI: **10.1002/jrs.2125**.

42 M. B. Seasholtz and B. Kowalski, The parsimony principle applied to multivariate calibration, *Anal. Chim. Acta*, 1993, **277**(2), 165–177, DOI: **10.1016/0003-2670(93)80430-S**.

43 S. Sapru, M. Vardhan, Q. Li, Y. Guo, X. Li and D. Saxena, E-cigarettes use in the United States: Reasons for use, perceptions, and effects on health, *BMC Public Health*, 2020, **20**(1), 1–10, DOI: **10.1186/s12889-020-09572-x**.

44 M. Baranska, J. C. Dobrowolski, A. Kaczor, K. Chruszcz-Lipska, K. Gorz and A. Rygula, Tobacco alkaloids analyzed by Raman spectroscopy and DFT calculations, *J. Raman Spectrosc.*, 2012, **43**(8), 1065–1073, DOI: **10.1002/jrs.3127**.

45 E. J. Z. Krüsemann, J. L. A. Pennings, J. W. J. M. Cremers, F. Bakker, S. Boesveldt and R. Talhout, GC–MS analysis of e-cigarette refill solutions: A comparison of flavoring composition between flavor categories, *J. Pharm. Biomed. Anal.*, 2020, **188**, DOI: **10.1016/j.jpba.2020.113364**.

46 A. El-Hellani, R. El-Hage, R. Baalbaki, *et al.*, Free-Base and Protonated Nicotine in Electronic Cigarette Liquids and Aerosols, *Chem. Res. Toxicol.*, 2015, **28**(8), 1532–1537, DOI: **10.1021/acs.chemrestox.5b00107**.

47 J. L. A. Pennings, A. Havermans, C. G. G. M. Pauwels, E. J. Z. Krüsemann, W. F. Visser and R. Talhout, Comprehensive

Dutch market data analysis shows that e-liquids with nicotine salts have both higher nicotine and flavour concentrations than those with free-base nicotine, *Tob. Control.*, 2023, **32**, e78–e82, DOI: **10.1136/tobaccocontrol-2021-056952**.

48 A. M. Leventhal, D. R. Madden, N. Peraza, *et al.*, Effect of Exposure to e-Cigarettes With Salt vs Free-Base Nicotine on the Appeal and Sensory Experience of Vaping: A Randomized Clinical Trial, *JAMA Network Open*, 2021, **4**(1), e2032757, DOI: **10.1001/jamanetworkopen.2020.32757**.