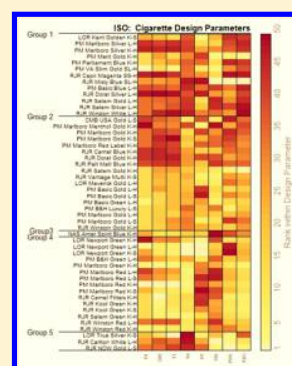


## Multivariate Statistical Analysis of Cigarette Design Feature Influence on ISO TNCO Yields

Kimberly A. Agnew-Heard,<sup>†</sup> Vicki A. Lancaster,<sup>†</sup> Roberto Bravo,<sup>‡</sup> Clifford Watson,<sup>‡</sup> Matthew J. Walters,<sup>\*,†</sup> and Matthew R. Holman<sup>†</sup><sup>†</sup>Center for Tobacco Products, U.S. Food and Drug Administration, 10903 New Hampshire Avenue, Silver Spring, Maryland 20993, United States<sup>‡</sup>National Center for Environmental Health, Centers For Disease Control and Prevention, 4770 Buford Highway, NE, Atlanta, Georgia 30341, United States

**ABSTRACT:** The aim of this study is to explore how differences in cigarette physical design parameters influence tar, nicotine, and carbon monoxide (TNCO) yields in mainstream smoke (MSS) using the International Organization of Standardization (ISO) smoking regimen. Standardized smoking methods were used to evaluate 50 U.S. domestic brand cigarettes and a reference cigarette representing a range of TNCO yields in MSS collected from linear smoking machines using a nonintense smoking regimen. Multivariate statistical methods were used to form clusters of cigarettes based on their ISO TNCO yields and then to explore the relationship between the ISO generated TNCO yields and the nine cigarette physical design parameters between and within each cluster simultaneously. The ISO generated TNCO yields in MSS are 1.1–17.0 mg tar/cigarette, 0.1–2.2 mg nicotine/cigarette, and 1.6–17.3 mg CO/cigarette. Cluster analysis divided the 51 cigarettes into five discrete clusters based on their ISO TNCO yields. No one physical parameter dominated across all clusters. Predicting ISO machine generated TNCO yields based on these nine physical design parameters is complex due to the correlation among and between the nine physical design parameters and TNCO yields. From these analyses, it is estimated that approximately 20% of the variability in the ISO generated TNCO yields comes from other parameters (e.g., filter material, filter type, inclusion of expanded or reconstituted tobacco, and tobacco blend composition, along with differences in tobacco leaf origin and stalk positions and added ingredients). A future article will examine the influence of these physical design parameters on TNCO yields under a Canadian Intense (CI) smoking regimen. Together, these papers will provide a more robust picture of the design features that contribute to TNCO exposure across the range of real world smoking patterns.



## INTRODUCTION

Many cigarette manufacturers routinely report machine smoke deliveries of tar, nicotine, and carbon monoxide (TNCO), which have been used to categorize, rank, and monitor cigarette brands in the United States.<sup>1–3</sup> Of particular significance, European Union regulations as early as 1992 mandated a limit on machine-measured tar for all cigarettes and in 2004 established a maximum delivery of 10 mg of tar, 1 mg of nicotine, and 10 mg of carbon monoxide (CO) using the International Organization for Standardization (ISO) protocol.<sup>4</sup> The new mandates spurred a change in cigarette design (i.e., increasing filter ventilation) and tobacco blend composition to significantly reduce ISO generated TNCO yields.<sup>5–7</sup> Data collected throughout the last few decades have proven that mainstream smoke (MSS) yields generated from nonintense smoking conditions (i.e., Federal Trade Commission and ISO) represent an estimated low end of exposure and does not directly correlate to the smoker's exposure.<sup>5</sup>

In 2012, the FDA published an established list of 93 harmful and potentially harmful constituents (HPHCs), including nicotine and carbon monoxide, in MSS and tobacco filler, as required under section 904(e) of the Federal Food, Drug, and Cosmetic Act.<sup>8,9</sup> As numerous regulatory agencies have

required the reporting and reduction in select constituents, tobacco manufacturers have moved toward designing cigarettes with lower TNCO yields under nonintense smoking parameters, accomplished by optimizing the physical design components, tobacco blend composition, and tobacco mass to meet regulatory requirements.<sup>3,4,9–13</sup> However, it can be difficult to relate cigarette design to TNCO yields because people do not smoke the way the machine “smokes” a cigarette. In addition, proprietary cigarette design parameters often are not reported in studies that evaluate commercial cigarette products. Consequently, research and benchmark studies have focused on a limited set of cigarette physical design parameters, such as filter ventilation, in addition to the cigarette's dimensions (circumference and lengths: rod, filter, and overwrap), draw resistance, pressure drop, paper porosity, and tobacco filler weight, which can be independently measured.

In evaluating TNCO and other constituents with ISO machine-measured mainstream cigarette smoke deliveries, filter ventilation is often a dominant variable to significantly alter or

Received: March 21, 2016

Published: May 25, 2016

reduce TNCO deliveries through the dilution of MSS. Drawbacks to increasing filter ventilation as a way to reduce actual human exposure (as opposed to machine-measured emissions) include smokers possibly compensating to attain an acceptable inhaled nicotine level and a satisfactory “mouth-feel” from the “heavier” or harsher taste. Smokers tend to take larger and more frequent puffs, block the ventilation holes, or switch to a less ventilated cigarette as compensatory behaviors when filter ventilation is increased.<sup>7,14</sup> Unfortunately, machine-measured smoking conditions only provide an estimate of smoke emissions under a fixed set of conditions (puff duration, volume, and frequency) and cannot mimic all human smoking behavior that can alter human exposure to MSS constituents. When evaluating the range of smokers’ exposure to HPHCs and tar in mainstream cigarette smoke, ISO and Canadian Intense (CI) can potentially represent the lower and higher ends of exposure, respectively.<sup>15</sup> Thus, the main utility of machine smoking is to compare MSS deliveries between brands, but also to investigate how cigarette design influences ISO generated TNCO delivery.

Prior tobacco literature generally evaluated the dependence of TNCO yields on cigarette physical design parameters using a sequential approach.<sup>15–21</sup> For example, the correlation coefficient between TNCO yields and physical design measurements are provided and used to describe how individual physical design measurements influence the amount of TNCO emitted by the cigarette. Other studies conducted bivariate or multivariate regressions with tar or nicotine on commercial cigarettes; however, most of these studies do not account for how individual physical design parameters impact each other while affecting TNCO deliveries.<sup>21–23</sup>

This article summarizes the results from a series of multivariate techniques to evaluate the relationships between TNCO yields under ISO smoking conditions and nine cigarette design parameters simultaneously. Multivariate analyses involved principal components (PCs) analysis, cluster analysis, and partial least squares regression (PLSR). The aim is to evaluate 50 cigarettes that represent a wide range of U.S. commercially marketed, filtered cigarettes from different manufacturers and diverse brand styles on the market in 2011. A wide range of TNCO yields measured under ISO smoking conditions and physical design properties, such as filter ventilation, paper porosity, draw resistance, pressure drop, tobacco weight, and cigarette dimensions (rod and filter lengths, circumference, and overwrap) were explored in the cigarettes for this study. Although the physical design properties and ISO generated TNCO yields were measured, proprietary tobacco composition and other cigarette material components, additives, and ingredients information were not considered in this study.

## ■ EXPERIMENTAL PROCEDURES

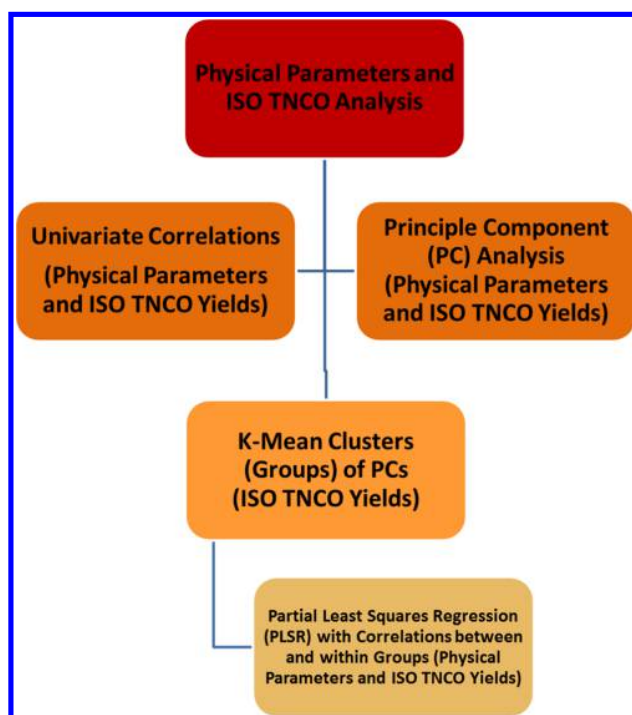
**Cigarette Brand Selection.** Fifty commercial filtered cigarettes were purchased in 2011 from different retail stores in metropolitan Atlanta, Georgia over a 1-month period. After purchase, the unopened cigarette packages were assigned an identification number and logged into a database developed using Paradox software (Borland Inprise, Scotts Valley, California). The cigarettes in their original packaging were sealed using plastic bags and stored at  $-70^{\circ}\text{C}$  until tested. Only authorized personnel had access to the samples. The 50 cigarette brand styles selected consisted of 35 brands having 53% of the total cigarette market share, as estimated from the 2007 Maxwell reports.<sup>24</sup> An additional 15 brands were selected to encompass brands with low market share, as well as different cigarette design features and ISO

machine-measured tar groupings. The chosen filtered brand styles include nine cigarettes with ISO machine-measured tar levels of 1–6 mg; 25 with 7–12 mg; and 16 with >13 mg of tar, and the cigarettes were labeled as king-size (K) or 100s (L) in hard (H) or soft (S) packaging. Fourteen cigarettes were identified as mentholated by label statements, package coloring, and manufacturer’s description. All cigarettes are manufactured by either Commonwealth Brands, Inc. (CMB); Reynold American Subsidiaries Lorillard, Inc. (LOR), Santa Fe Natural Tobacco Company, LLC. (NAS), and R.J. Reynolds Tobacco Company (RJR); or Philip Morris USA. The selection of these products should be considered a convenience sample and not a nationally representative sampling of marketed products.

**TNCO Smoke Analysis.** The tar, nicotine, and carbon monoxide (TNCO) yields measured in the mainstream smoke were determined as discussed below. Kentucky 3R4F research cigarettes (University of Kentucky Tobacco and Health Research Institute, Lexington, Kentucky) were analyzed with the commercial cigarettes for quality control. Prior to smoking, the cigarettes and Cambridge filter pads (CFPs) were conditioned at  $24^{\circ}\text{C}$  and 60% relative humidity for at least 48 h in an environmental chamber per ISO 3402:1999.<sup>25</sup> Cigarettes were smoked with a single 44 mm diameter glass fiber CFP (Fidus Instrument Corporation, Richmond, Virginia) using an automated SM450 Cerulean 20-port linear smoking machine (Cerulean, Milton Keynes, UK) equipped with a CO analyzer. Twenty replicates of five cigarettes per brand were smoked per ISO 3308:2000 to a butt length of 23 mm or the length of the filter overwrap plus 3 mm, whichever was longer.<sup>26</sup> The gas phase portion of the cigarette mainstream smoke was collected in vapor phase collection bags, and the percentage by volume of CO (%CO) was determined using a Filtrona ATCOM 302 nondispersive IR analyzer (Filtrona Instruments & Automation Ltd., Milton Keynes, United Kingdom) in accordance with ISO 8454:2009.<sup>27</sup> The total particle matter was determined following ISO 4387:2008 by calculating the weight difference in the CFP before and after the smoking process. ISO 10315:2000 and ISO 10362-1:1999 were implemented to determine nicotine and water concentrations in mainstream smoke, respectively. Nicotine and water were extracted from each CFP with 2-propanol and analyzed by a 6890 gas chromatograph with flame ionization and a thermal conductivity detector, respectively (Agilent Technologies, Alpharetta, Georgia).<sup>28–30</sup> The tar content was calculated by subtracting the water and nicotine content from the total particulate matter (TPM). Twenty replicates were smoked, extracted, and analyzed for all cigarette brand styles and 3R4F reference cigarette.

**Physical Parameters.** After conditioning the cigarette, seven replicates of physical dimensions, rod length (RL), filter length (FL), circumference (CIR), and overwrap (OW), along with draw resistance (PDO), pressure drop (PDS), and filter tip ventilation (TV) were measured using C2 autosampler equipment (Cerulean, Milton Keynes, UK), which contains an independently established charge coupled device technology and is compatible with ISO 2971:2013 and ISO 6565:2011.<sup>31,32</sup> The methods ISO 2965:2009 and CORESTA Recommended Method No. 40 were used with a PPM1000 M instrument (Cerulean, Milton Keynes, UK) to determine air permeability and paper porosity (PP).<sup>33,34</sup> Seven replicates of cigarette tobacco filler weights (TW) were measured manually for all cigarettes.

**Statistical Analysis.** All analyses were conducted in R, a programming language and software environment for statistical computing and graphics, freely available under the GNU General Public License. Multivariate analyses involved first transforming the correlated ISO TNCO yields from the 51 cigarettes into a set of uncorrelated principal components and then using the principal components (PCs) in a K-means clustering algorithm to form groups of “like” cigarettes. Once the clusters were determined, the relationship between the original ISO TNCO yields and the nine physical design parameters within each group using partial least squares regression (PLSR) was explored. A flow diagram in Figure 1 illustrates an overview of the statistical analysis, and a more detailed explanation is provided below.



**Figure 1.** Flow diagram for statistical analysis of the physical design parameters and ISO generated TNCO yields of 51 cigarettes.

To evaluate the relationship between the ISO generated TNCO yields and physical design parameters simultaneously, our initial step was to transform the TNCO yields into a set of linearly uncorrelated variables called principal components (PCs). The transformation is defined so that the first principal component has the largest possible variance (i.e., accounts for as much of the variability in the data as possible), and each succeeding component, in turn, has the highest variance possible under the constraint orthogonal (uncorrelated) to the preceding components. All three principal components for tar, nicotine, and carbon monoxide are retained to prevent loss of information.

The second step was to identify clusters among the 51 cigarettes in this study using the transformed TNCO emissions and then to describe the TNCO emissions under ISO conditions within the clusters using the physical design parameters. Since there were no established groups among the TNCO emissions or physical parameters that could be used to define the clusters, an unsupervised learning technique was implemented. Clustering algorithms in the unsupervised learning finds the hidden structure in data that has no prior grouping. These clustering algorithms are tools used to summarize a set of data with meaningful but unknown structure; since there is no direct measure of success, the validity of particular clustering algorithm is based on heuristic arguments. The approach employed with these data, combined two common multivariate methods: principal component analysis and discriminant analysis. These combined methods often produce clusters that are well-defined and interpretable. The algorithm (adeqnet R Package, version: 1.4–2, 2014–05–13, GPL (≥2)) first transformed the TNCO emissions into their principal components and then used a K-means clustering algorithm to define the clusters. The optimal number of clusters was determined by sequentially increasing the number of clusters followed by calculating the Bayesian Information Criterion, which is based on the log likelihood. Once the number of clusters was determined, a discriminant analysis was used to evaluate the cluster assignments using a probabilistic assessment. Using the principal components as the inputs to discriminant analysis avoids the multicollinearity problem posed by the TNCO emissions (Tar versus Nic,  $r^2 = 0.93$ ; Tar versus CO,  $r^2 = 0.91$ ; and Nic versus CO,  $r^2 = 0.82$ ).

Once the clusters were defined, the relationships among the TNCO emissions under ISO conditions and physical design parameters within and between each cluster were evaluated using partial least squares regression (PLSR). The PLSR approach constructs predictive models in cases where the data are highly correlated, the number of predictor variables may be greater than the number of observations, and the number of variables on both sides of the equation (predictor and response) is greater than one. Orthogonal components defined by PLSR capture most of the variability in the predictor and response variables while maximizing the correlation between these two matrices. The components are calculated sequentially from the covariance structure between the response and predictor variables. The first component is constructed to maximize the correlation between the response and predictor variables, then the second component is constructed on the deflated covariance matrix; the variability accounted for the first component has been removed. This continues until the number of components equals the number of predictor variables. This estimation approach provides components that are independent from one another. The number of components that remain in the model is a function of the variability explained by each component. Since the original variables are usually correlated and contain redundant information, the number of components will be less than the number of the original variables. Unlike multiple linear regression where the response and predictor variables are linked by a single set of coefficients and residual error, partial least-squares regression decomposes the response and predictor variables simultaneously as a product of orthogonal components, where each component comprises a set of scores, loadings, and residuals. The constraint on the components is that they account for as much of the covariance between the response and predictors as possible. Loadings indicate the contribution of a particular variable to the component, scores indicate how much a particular observation contributes to the component, and the residuals reflect what is not explained by the model.

## RESULTS

**Descriptive Statistics.** The means of the 20 TNCO and seven physical parameter replicates are displayed in Table 1. TNCO yields per cigarette in MSS ranged from 1.12–17.02 mg/cig tar, 0.13–2.15 mg/cig nicotine, and 1.60–17.26 mg/cig CO when measured under ISO smoking conditions. King-size cigarette rod lengths (RL) are 79.14 to 84.00 mm, while 100s ranged from 97.00 to 99.29 mm. The range in filter length (FL) is between 15.86 mm and 32.00 mm with an overwrap length (OW) that is typically 4.5 mm greater than the filter length. An average circumference of 24.4 mm is measured for 51 cigarettes where the superslim is 17.1 mm, and the largest is 27.9 mm (3R4F). Filter ventilation (TV) ranged from 0.14% (non-ventilated) to 66.60% (highly ventilated) with a mean draw resistance (PDO) of 106.01 mmWG, inclusive of 7.73 mmWG measured for cigarette 33 and paper porosity (PP) of 45.49 CORESTA. The largest ranges in physical design measurements are seen in tobacco weight (TW), 401.03 to 880.71 mg, and pressure drop (PDS), 10.06 to 193.96 mmWG. Draw resistance across the cigarette rod ranged from 7.73 to 134.69 mmWG, resulting in a calculated difference between pressure drop and draw resistance of –0.21 to 89.43 mmWG. Hard and soft cigarette packaging comparisons between the same brand styles revealed differences in length and tobacco weight for cigarette pairs 5 and 6 and pairs 23 and 24.

**Univariate Correlations.** Univariate correlations among the TNCO yields and the nine cigarette physical design parameters (RL, FL, CIR, OW, PP, TW, TV, PDO, and PDS) are listed in Table 2. Only two of the physical design parameters, paper porosity and circumference ( $r < |0.25|$ ), are not significantly correlated ( $p$ -value  $< 0.05$ ) with any of the

Table 1. ISO Generated TNCO Yields and Physical Design Parameters of 51 Cigarettes Grouped by K-Mean Clusters

ID	cigarette brand style	TAR <sup>a</sup>	NIC <sup>a</sup>	CO <sup>a</sup>	RL <sup>b</sup>	OW <sup>b</sup>	FL <sup>b</sup>	Cir <sup>b</sup>	TV <sup>c</sup>	PP <sup>d</sup>	TW <sup>e</sup>	PDO <sup>f</sup>	PDS <sup>f</sup>	cluster
13	PM B&H Green L-H	13.26	1.13	13.22	98.00	31.86	26.71	24.8	26.70	34.87	714.57	105.90	132.49	1
48	RJR Winston Red L-H	13.53	1.37	14.19	98.00	35.00	30.86	24.4	19.13	49.99	735.66	105.64	125.50	1
5	LOR Newport Green K-H	13.79	1.02	15.67	80.14	25.29	20.71	24.8	0.14	15.90	635.00	132.56	132.46	1
20	PM Marlboro Green K-H	13.99	1.09	12.09	82.86	24.71	20.29	24.9	15.16	50.40	657.59	94.50	105.59	1
24	PM Marlboro Red K-S	13.99	1.09	13.45	83.43	25.43	20.86	24.7	14.64	72.04	720.94	106.31	118.84	1
21	PM Marlboro Red L-H	14.08	1.18	14.24	98.43	32.14	27.00	24.7	12.81	51.40	777.66	106.99	119.37	1
23	PM Marlboro Red K-H	14.10	1.11	13.09	79.57	24.00	18.43	24.8	8.86	53.64	662.91	105.06	111.64	1
22	PM Marlboro Red L-S	14.29	1.21	13.63	98.00	32.00	26.86	24.9	13.99	57.77	790.37	107.09	120.19	1
49	RJR Winston Red K-H	14.59	1.35	16.61	82.86	25.14	20.86	24.4	17.16	41.17	698.97	114.91	129.47	1
32	RJR Camel Filters K-H	14.63	1.27	12.70	83.00	24.14	21.00	24.4	21.04	56.46	715.69	91.26	106.54	1
44	RJR Salem Green K-H	14.90	1.22	14.65	83.00	25.00	20.43	24.5	13.07	38.36	692.31	100.24	110.51	1
4	LOR Newport Green L-H	14.92	1.27	13.79	99.00	33.43	28.57	24.7	21.07	22.36	723.86	99.41	120.69	1
37	RJR Kool Green K-H	15.27	1.30	13.90	83.00	24.86	21.00	24.3	19.97	57.23	659.91	94.57	109.53	1
6	LOR Newport Green K-S	15.77	1.16	17.26	84.00	25.43	20.57	24.8	0.16	17.30	648.80	136.23	136.01	1
38	RJR Kool Green K-S	16.06	1.27	14.37	83.00	25.00	20.86	24.5	20.39	50.27	671.31	94.47	108.94	1
50	RJR Winston White L-H	4.67	0.53	8.00	98.00	36.00	31.86	24.4	58.93	50.53	712.29	111.66	193.96	2
36	RJR Doral Silver L-H	4.93	0.46	7.94	98.14	35.00	30.29	24.4	51.60	51.49	729.43	115.29	182.40	2
45	RJR Salem Silver L-H	5.15	0.49	8.64	98.00	34.43	30.86	24.3	50.07	43.57	767.29	120.89	191.56	2
9	PM Basic Blue L-H	5.35	0.49	8.32	98.71	35.71	31.29	24.8	43.74	37.40	628.83	120.13	175.44	2
27	PM Marlboro Silver K-H	5.43	0.51	6.90	83.00	32.00	25.86	24.8	33.87	55.14	584.73	73.43	97.24	2
26	PM Marlboro Silver L-H	5.94	0.56	7.99	98.86	36.00	31.43	24.7	45.06	48.74	683.26	118.91	174.93	2
2	LOR Kent Golden K-S	6.64	0.68	8.04	83.71	31.57	24.29	24.8	41.26	26.49	657.03	121.94	170.89	2
28	PM Merit Gold K-H	6.92	0.64	8.81	83.00	31.57	27.00	24.5	38.47	51.79	619.06	127.26	173.53	2
42	RJR Salem Gold L-H	7.30	0.66	10.29	98.00	34.71	30.71	24.3	46.61	37.17	767.57	113.40	175.16	2
30	PM VA Slim Gold SL-H	7.51	0.71	8.72	98.86	36.43	31.29	23.0	36.66	49.83	623.96	134.69	186.40	2
29	PM Parliament Blue K-H	8.09	0.71	9.55	79.14	29.86	20.00	24.8	36.73	34.17	587.00	112.01	156.99	2
39	RJR Misty Blue SL-H	8.39	0.75	8.89	99.00	37.00	31.86	22.9	43.61	60.03	672.01	109.21	161.33	2
33	RJR Capri Magenta SS-H	8.72	0.86	5.76	97.71	32.00	27.43	17.1	60.59	70.66	401.03	7.73	10.06	2
8	NAS Amer Spirit Blue K-H	17.02	2.15	14.90	83.86	27.14	23.14	24.2	21.89	42.39	880.71	120.10	143.83	3
34	RJR Carlton White L-H	1.12	0.13	1.60	97.00	35.00	30.00	24.5	62.31	47.44	584.71	61.59	114.84	4
40	RJR NOW Gold L-S	2.03	0.25	3.08	98.00	34.71	31.00	24.4	55.09	52.27	588.96	60.73	105.11	4
7	LOR True Silver K-S	3.93	0.45	4.32	83.29	29.71	15.86	24.3	66.60	32.79	645.46	65.34	154.77	4
51	Kentucky Reference 3R4F	8.00	0.73	10.35	83.92	5.36	26.72	27.85	38.56	30.61	821.70	131.14	182.31	5
3	LOR Maverick Gold L-H	8.66	0.81	9.55	99.29	33.71	27.57	24.7	43.73	30.31	795.83	110.23	161.11	5
43	RJR Salem Gold K-H	9.12	0.82	10.99	83.00	31.14	26.29	24.4	32.53	44.70	636.89	105.91	137.89	5
16	PM Marlboro Gold L-S	9.41	0.81	10.94	98.29	35.86	31.29	24.9	29.44	44.10	730.07	116.87	150.43	5
41	RJR Pall Mall Blue K-H	9.51	0.92	9.28	83.00	31.00	26.86	24.4	37.47	45.34	703.53	88.86	126.64	5
10	PM Basic Gold L-H	9.69	0.78	13.03	98.00	35.71	30.86	24.7	20.49	55.03	682.36	126.49	151.31	5
47	RJR Winston Gold K-H	9.73	1.01	11.37	83.00	31.00	27.00	24.3	33.94	46.14	639.14	105.23	139.91	5
12	PM Basic Green L-H	9.73	0.78	12.38	98.00	35.86	31.00	24.8	21.66	53.00	714.19	119.37	142.89	5
46	RJR Vantage Multi K-S	9.92	0.84	10.46	83.00	31.29	27.14	24.2	34.90	34.16	610.14	121.54	149.33	5
11	PM Basic Gold L-S	9.93	0.82	13.42	98.00	36.00	32.00	24.7	23.47	45.36	712.29	121.39	147.70	5
35	RJR Doral Gold K-H	9.97	0.89	11.60	82.86	30.86	26.86	24.4	24.14	49.01	592.27	102.67	124.09	5

Table 1. continued

ID	cigarette brand style	TAR <sup>a</sup>	NIC <sup>a</sup>	CO <sup>a</sup>	RL <sup>b</sup>	OW <sup>b</sup>	FL <sup>b</sup>	CL <sup>b</sup>	TV <sup>c</sup>	PP <sup>d</sup>	TW <sup>e</sup>	PDO <sup>f</sup>	PDS <sup>f</sup>	cluster
31	RJR Camel Blue K-H	10.08	0.90	11.00	83.00	30.86	26.86	24.3	25.49	64.60	608.00	113.56	137.74	5
14	PM B&H Luxury L-S	10.13	0.91	12.33	98.71	36.00	31.14	24.8	28.44	42.23	637.10	116.51	148.39	5
19	PM Marlboro Gold K-S	10.17	0.83	10.90	83.00	31.71	26.86	24.8	23.17	35.61	595.60	115.73	138.43	5
15	PM Marlboro Gold L-H	10.19	0.94	12.19	98.71	36.00	31.43	24.7	30.11	54.79	716.19	111.30	144.46	5
18	PM Marlboro Gold K-H	10.25	0.83	11.00	83.00	31.86	26.86	24.8	25.20	32.83	591.86	111.64	135.91	5
17	PM Marlboro Menthol Gold K-H	10.38	0.85	9.95	83.00	32.00	27.00	24.7	22.11	41.04	607.53	68.54	82.96	5
1	CMB USA Gold L-S	10.63	0.93	12.66	98.43	34.86	29.43	24.3	18.89	54.30	599.90	128.03	149.50	5
25	PM Marlboro Red Label K-H	11.32	0.92	11.27	83.00	32.00	26.86	24.8	23.93	55.83	627.63	101.90	124.20	5

<sup>a</sup>mg/cig. <sup>b</sup>mm. <sup>c</sup>%. <sup>d</sup>CORESTA. <sup>e</sup>mg. <sup>f</sup>mmWG.

TNCO yields. Pressure drop is not significantly correlated with tar and nicotine,  $r \approx 0.2$  ( $p$ -value  $> 0.05$ ), but significantly correlated with CO,  $r \approx 0.5$  ( $p$ -value  $< 0.05$ ). The reverse is true for draw resistance, where tar and nicotine are significantly correlated,  $r \approx -0.3$  ( $p$ -value  $< 0.05$ ), but the correlation coefficient for CO is close to 0. For the remainder of the physical parameters, the trends for TNCO yields are similar; all three exhibit a weak positive correlation with tobacco weight ( $r \approx 0.3$ ,  $p$ -value  $< 0.05$ ), a strong negative correlation with filter ventilation ( $r \approx -0.7$ ,  $p$ -value  $< 0.05$ ), and a moderate negative correlation with the lengths (i.e., rod length, overwrap, and filter length) ( $0.5 < r < -0.2$ ).

**ISO TNCO Group Identification.** Figure 2 displays the results of the K-means clustering algorithm where the clusters are characterized by their ISO TNCO emissions. Five clusters were selected based on the Bayesian Information Criterion and the probabilistic assessment of the discriminant analyses shown in Figure 3. Group 3 contains a single cigarette with the highest ISO tar and nicotine emissions, and Group 4 contains three cigarettes with the lowest. Groups 3 and 4 will not be discussed further due to the limited number of cigarettes within these clusters.

Clusters 1, 2, and 5 were used to evaluate the relationship between ISO generated TNCO emissions and the physical design parameters. Group 1 contains 15 cigarettes, comprising the highest ISO emissions of tar (13.26–16.06 mg/cig), nicotine (1.02–1.37 mg/cig), and CO (12.09–17.26 mg/cig) of the three remaining clusters. For these cigarettes, the lowest probability of group membership is  $>0.99$  for cigarette 13 (Figure 3). Group 2 contains 13 cigarettes comprising the lowest ISO emissions of tar (4.67–8.72 mg/cig), nicotine (0.46–0.86 mg/cig), and CO (5.76–10.29 mg/cig). The two lowest probabilities of group membership are 0.51 for cigarette 39 and 0.56 for cigarette 29, and the remaining 11 cigarettes all have probabilities of group membership  $>0.85$ . Group 5 comprises 19 cigarettes with TNCO emissions that are between the range of Groups 1 and 2 (see Table 1). The lowest probability of group membership is 0.63 for Reference 3R4F; the remaining 18 cigarettes all have probabilities of group membership  $>0.83$ .

Group 2 encompasses the widest range of ISO generated tar yield cigarettes with a 40% probability of containing the Kentucky research cigarette 3R4F (cigarette 51) and 50–55% probability of cigarette 29 and cigarette 39 belonging to this cluster. Similarly, there is a 60% probability that 3R4F belongs to Group 5, along with 45–50% chance of it containing cigarette 29 and cigarette 39. The ISO tar yields of these three cigarettes lie between Groups 2 and 5, approximately 8 mg/cig, which seems to be on the boundary between these two groups. For all cigarettes within this study, cigarette 29 consists of the shortest rod length, cigarette 39 has the longest overwrap, and 3R4F is the cigarette with the shortest overwrap and largest circumference.

**Partial Least Squares Regression.** The relationship between the ISO generated TNCO emissions and the physical parameters was explored using partial least squares regression (PLSR), and a separate PLSR was run for each group (Figures 2 and 3). The maximum number of components available to describe the relationship between the ISO TNCO deliveries and physical parameters is nine, the number of physical parameters. When all nine components were retained in the three models, 100% of the variability from the physical parameters is explained and approximately 82% of the

Table 2. Univariate Correlations for the ISO Generated TNCO Yields and Physical Design Parameters<sup>a</sup>

parameter	TAR (mg/cig)	NIC (mg/cig)	CO (mg/cig)	RL (mm)	OW (mm)	FL (mm)	CIR (mm)	TV (%)	PP (CORESTA)	TW (mg)	PDO (mmWG)	PDS (mmWG)
TAR (mg/cig)	1.00	<b><u>0.93</u></b>	<b><u>0.91</u></b>	<b><u>-0.32</u></b>	<b><u>-0.42</u></b>	<b><u>-0.49</u></b>	0.07	<b><u>-0.86</u></b>	-0.04	<b><u>0.30</u></b>	0.19	<b><u>-0.38</u></b>
NIC (mg/cig)		1.00	<b><u>0.82</u></b>	<b><u>-0.28</u></b>	<b><u>-0.37</u></b>	<b><u>-0.42</u></b>	0.00	<b><u>-0.71</u></b>	-0.01	<b><u>0.39</u></b>	0.16	<b><u>-0.32</u></b>
CO (mg/cig)			1.00	-0.22	<b><u>-0.33</u></b>	<b><u>-0.32</u></b>	0.25	<b><u>-0.90</u></b>	-0.15	<b><u>0.39</u></b>	<b><u>0.49</u></b>	-0.09
RL (mm)				1.00	<b><u>0.65</u></b>	<b><u>0.79</u></b>	-0.21	<b><u>0.35</u></b>	0.17	0.21	0.02	0.23
OW (mm)					1.00	<b><u>0.68</u></b>	<b><u>-0.33</u></b>	<b><u>0.34</u></b>	0.19	-0.18	-0.04	0.17
FL (mm)						1.00	-0.09	<b><u>0.38</u></b>	0.19	0.06	0.14	<b><u>0.34</u></b>
CIR (mm)							1.00	<b><u>-0.30</u></b>	<b><u>-0.40</u></b>	<b><u>0.53</u></b>	<b><u>0.58</u></b>	<b><u>0.49</u></b>
TV (%)								1.00	0.09	-0.19	<b><u>-0.39</u></b>	<b><u>0.26</u></b>
PP (CORESTA)									1.00	-0.16	<b><u>-0.35</u></b>	<b><u>-0.32</u></b>
TW (mg)										1.00	<b><u>0.48</u></b>	<b><u>0.45</u></b>
PDO (mmWG)											1.00	<b><u>0.76</u></b>
PDS (mmWG)												1.00

<sup>a</sup>Bold faced and underlined correlation coefficients are statistically different from 0 at  $p$ -value <0.05.

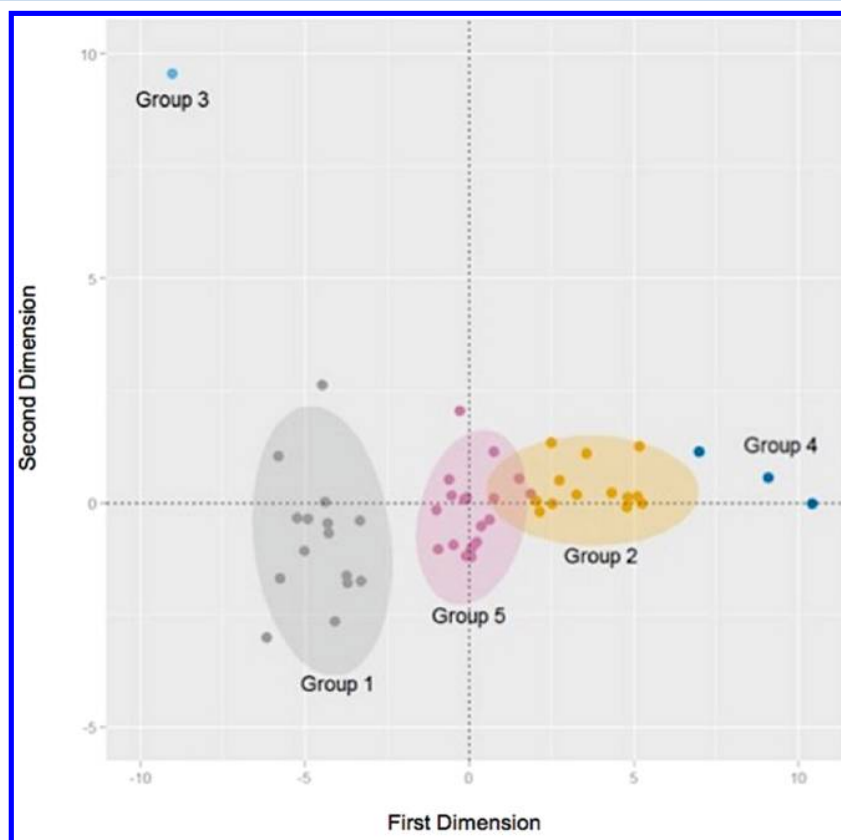
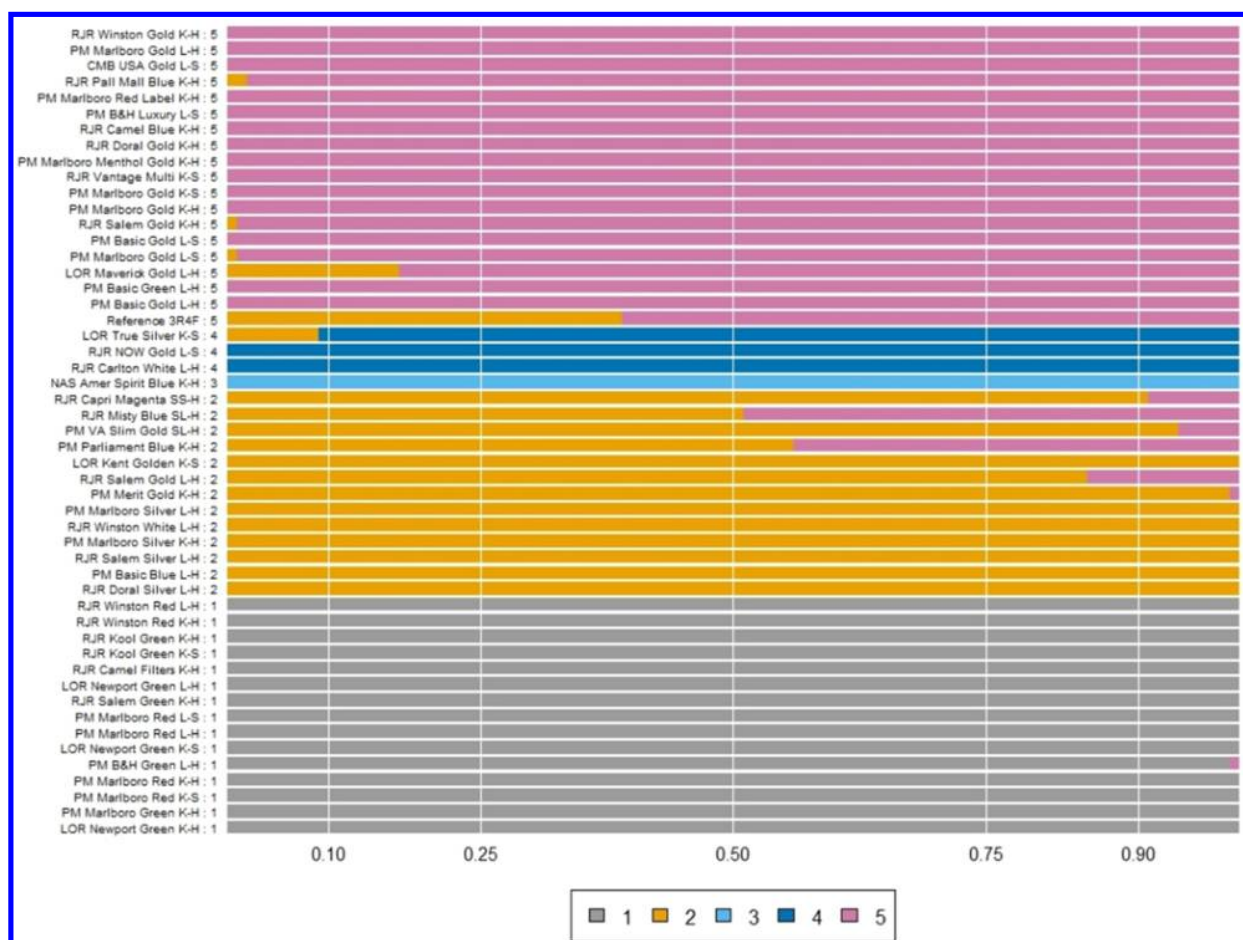


Figure 2. K-means cluster (group) identification using ISO generated TNCO yields: Group 1, 13.26–16.06 mg/cigarette tar (gray); Group 2, 4.67–8.72 mg/cigarette tar (orange); Group 3, 17.02 mg/cigarette tar (light blue); Group 4, 1.12–3.93 mg/cigarette tar (dark blue); and Group 5, 8.00–11.32 mg/cigarette tar (pink).

variability in the ISO TNCO emissions. One can infer that approximately 18% of the variability in the ISO TNCO emissions is not accounted for by these physical parameters. Three components were retained for each of the three clusters; for Group 1, this accounted for 87% of the cumulative variability in the physical parameters and 61% of the cumulative variability in the ISO TNCO emissions; for Group 2, 89% physical and 65% ISO TNCO; and for Group 5, 81% physical and 65% ISO TNCO (Table 3). The correlations between the

emission and physical scores for each component can also provide information on the number of components to retain. Table 3 displays the correlations between the scores for each group by component combination. The correlation with component 1 for each of the three groups is >0.82.

The loadings for each component shed light on how the relationships between ISO TNCO emissions from machine smoking and physical design features can vary between the three groups. Table 4 contains the loadings for the first two



**Figure 3.** Probability of cluster (group) membership of each cigarette determined by ISO generated TNCO yields, where the length of the bar depicts the probability of membership in Group 1 (gray), Group 2 (orange), Group 3 (light blue), Group 4 (dark blue), and Group 5 (pink).

**Table 3. Summary of the ISO Generated TNCO and Physical Design Parameter Variations Explained by PLSR for Groups 1, 2, and 5 by PLSR Components<sup>a</sup>**

group	ISO TNCO variation explained by components			physical variation explained by components			correlation ISO TNCO <sub>SCORES</sub> vs physical <sub>SCORES</sub>		
	1	2	3	1	2	3	1	2	3
1	0.251	0.146	0.213	0.389	0.311	0.175	0.923	0.739	0.678
2	0.385	0.068	0.197	0.471	0.354	0.067	0.880	0.353	0.652
5	0.372	0.211	0.067	0.379	0.321	0.115	0.822	0.837	0.735

<sup>a</sup>Calculated correlations between the ISO generated TNCO yields and physical design parameters predicted from the PLSR model.

components by group. Loadings within a component that are within  $\pm 10\%$  of the largest loading are bold faced and underlined. The loadings suggest that the physical parameters correlate with the ISO CO yield for Groups 1, 2, and 5, but the influence on ISO tar and nicotine yields varies by group. For example, the tar loading for Group 1 for both components is the smallest among the three loadings within a component. In fact for component 1, the tar loading for Group 1 is close to zero, an indication the physical parameters do not account for the tar variability cigarettes with  $>13$  mg tar/cig yielding group. To evaluate this further, Table 5 provides the correlations between the predicted ISO TNCO emissions from the PLSR models and the observed yields. For Group 1, the lowest correlation is for tar, 0.487, which confirms that the PLSR model does not fully explain the variability in tar for Group 1.

In contrast, for Group 2 the correlation is 0.814, and for Group 5, it is 0.857.

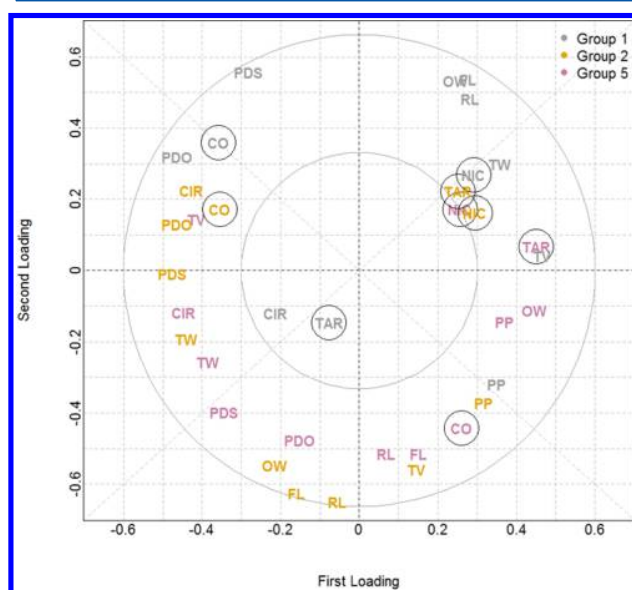
**Comparisons between Groups.** The ISO TNCO and physical design parameter loadings for the first two components for Groups 1, 2, 4 and 5 are plotted in Figure 4. The loadings for each group are identified by color, and the ISO TNCO loadings are distinguished from the physical design parameter loadings by enclosing the name of the TNCO loading within an open circle. The circles within the figure are provided to help with visual interpretation. Loadings near the circumference of the outer circle are more dominant than those near the center (0, 0). The proximity of the loadings to one another provides information on their correlation. Loadings that are close to one another are positively correlated (for example, OW, FL, and RL for Group 1); loadings that are in opposite direction from one another are negatively correlated (for example, PP and TV for Group 1).

**Table 4. Partial Least Squares Regression Loadings of the First and Second Components for ISO Generated TNCO Yields and Physical Design Features**

parameters		first component			second component		
		Group 1	Group 2	Group 5	Group 1	Group 2	Group 5
ISO TNCO	TAR	−0.077	0.252	<u>0.450</u>	−0.145	<u>0.223</u>	0.066
	NIC	0.290	0.294	0.257	0.269	0.162	0.171
	CO	<u>−0.359</u>	<u>−0.355</u>	0.259	<u>0.359</u>	0.173	<u>−0.441</u>
physical design	RL	0.281	−0.055	0.067	0.483	<u>−0.649</u>	<u>−0.514</u>
	OW	0.246	−0.214	<u>0.446</u>	<u>0.532</u>	−0.547	−0.113
	FL	0.277	−0.161	0.150	<u>0.537</u>	<u>−0.625</u>	<u>−0.515</u>
	TV	<u>0.464</u>	0.145	<u>−0.416</u>	0.042	−0.559	0.144
	PP	0.350	0.317	0.369	−0.319	−0.371	−0.143
	TW	0.356	<u>−0.440</u>	−0.386	0.298	−0.192	−0.256
	PDO	<u>−0.463</u>	<u>−0.463</u>	−0.150	0.319	0.131	−0.477
	PDS	−0.282	<u>−0.475</u>	−0.344	<u>0.556</u>	−0.008	−0.398
	CIR	−0.213	<u>−0.426</u>	<u>−0.446</u>	−0.120	0.224	−0.117

**Table 5. Correlations between the Predicted and Observed ISO Generated TNCO Yields from the PLSR Models**

ISO TNCO	correlation predicted vs observed		
	Group 1	Group 2	Group 5
TAR	0.487	0.814	0.857
NIC	0.878	0.815	0.632
CO	0.906	0.788	0.902

**Figure 4.** Loadings from the first and second components in Table 4 for Groups 1, 2, and 5 for ISO generated TNCO (identified with open circles) and physical design parameters. The proximity of the loadings in the two-dimensional space illustrates the association (the closer the loadings the more positively associated) between ISO generated TNCO and physical design parameters.

Group 5); and loadings that are orthogonal to one another are uncorrelated (PDS and RL for Group 2).

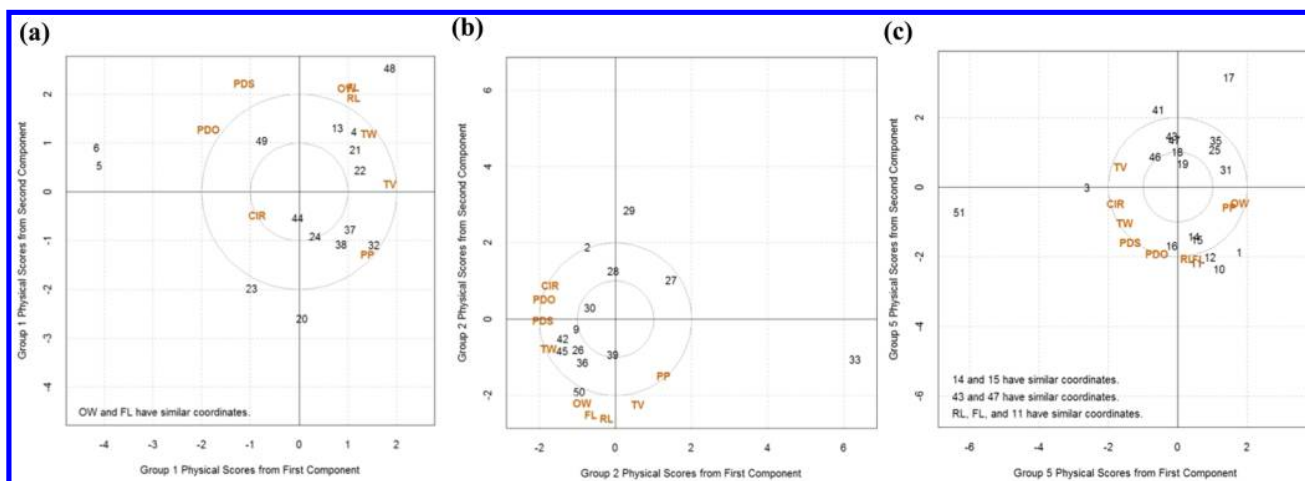
The most obvious difference between the three groups is with the placement of the ISO TNCO loadings. For Groups 2 and 5, the loadings for tar and nicotine are in close proximity to one another, an indication that within these groups tar and nicotine are correlated. For Group 1, the loadings for tar and nicotine are in different quadrants indicating they are less

correlated; the tar loading close to the center indicates that tar is not well represented in either component. It is also interesting to note that for Groups 2 and 5, nicotine is negatively correlated with tobacco weight but that for Group 1 it is positively correlated.

**Comparisons within Groups.** The 15 cigarette brand styles that encompass the highest ISO generated TNCO yields, with the exception of cigarette 8 in Group 3, are assembled in Group 1. Clustering of the kings and 100s cigarettes in this group are distinct, as demonstrated in Figure 5a where the king-size cigarettes (5, 6, 20, 23, 24, 32, 37, 38, 44, and 49) have a negative score on the first or second component or in the case of cigarette 23, a negative score on both components. Strong relationships are shown between cigarettes 32, 37, and 38 and paper porosity; between cigarettes 5, 6, and 49 and pressure drop and draw resistance; and for cigarette 23 and circumference. The 100s length brand styles (4, 13, 21, 22, and 48) have positive scores for both the first and second components and are correlated with lengths, tobacco weight, and filter ventilation. Cigarettes 5 and 6 appear to be outliers in Group 1 since these same mentholated cigarette brand styles are negatively correlated with filter ventilation. Cigarette 6 is encased in a soft package, has a slightly longer rod length, and higher tobacco weight, paper porosity, draw resistance, and pressure drop, resulting in higher ISO generated TNCO yield than cigarette 5, which is the same brand style in a hard package.

Figure 5b displays the scores for the first and second components in Group 2, and it divides the space into two groups, where the lowest tobacco filler weight brand styles (27, 29, and 33) have positive scores on the first component and are negatively correlated with tobacco weight. These cigarettes possess the smallest difference between draw resistance and pressure drop, where cigarette 29 exhibits the shortest lengths, and cigarette 33 has the highest tar, nicotine, filter ventilation, and paper porosity, and the smallest tobacco weight, circumference, pressure drop, and draw resistance. With the exception of cigarette brand style 30, all king-size length cigarettes (2, 27, 28, and 29) have positive scores for the second component and are negatively correlated with the length parameters.

Illustrated in Figure 5c are the 18 low 1–6 mg ISO generated tar cigarette brand styles in Group 5 where a clear distinction is observed between a the second component positive scores for the kings (17, 18, 19, 25, 31, 35, 41, 43, 46, and 47) and a negative scores for the 100s (1, 3, 10, 11, 12, 14, 15, and 16). In



**Figure 5.** Cigarette scores for the first and second components by (a) Group 1, (b) Group 2, and (c) Group 5, along with loadings for the first and second components for the physical design parameters. The proximity in the two-dimensional space illustrates the association (the closer the more positively associated) to the physical design parameters.

Group 5, cigarette 51 (3R4F research cigarette) is unique in that many of its physical parameters lie outside the range of medium 7–12 mg ISO tar yielding commercial cigarettes; this is reflected by the negative scores for both the first and second components. This reference cigarette has the lowest ISO machine measured tar and nicotine yields, paper porosity with the shortest overwrap, and filter length, along with the largest circumference, tobacco weight, draw resistance, and pressure drop. The longest filter lengths (>30 mm) and rod lengths (>98 mm) were as measured for cigarettes 10, 11, 12, 14, 15, and 16, which display a strong correlation with the filter and rod length loadings in Figure 5c.

Although Groups 3 and 4 are discrete clusters, the number of cigarette brand styles encompassed within these clusters is limited. Filler mass of the king-size cigarette brand style 8 is nearly 100 mg more than the largest 100s cigarette causing the tar and nicotine yields to lie outside the typical range of U.S. cigarettes within this study. Consequently, higher filler mass increases particulate phase emission, such as tar and nicotine. Although cigarette brand style 8 exhibits higher tar and nicotine than Group 1, its CO emission is comparable to cigarettes within Groups 1 and 5. Higher filter ventilation may contribute to the lower CO emission when comparing the Group 3 cigarette to Group 1 cigarettes under ISO smoking conditions.

The lowest measured ISO TNCO deliveries and highest filter ventilations in this study reside in Group 4 with three cigarette brand styles (7, 34, and 40). Only cigarette 33 in Group 2 possesses a lower draw resistance than Group 4 cigarettes. Cigarette 7 differs from the other two cigarettes in that it is king-sized, has lower paper porosity, higher pressure drop, and larger tobacco weight; whereas, the tobacco mass is approximately 50 mg more than cigarettes that are at least 14 mm longer. One distinctive feature of cigarette 7 is that the filter comprises a 15 mm conventional filter and an additional 10 mm plastic spacer for a total filter length of 25 mm.

## DISCUSSION

Under ISO smoking conditions, cigarette MSS constituent deliveries for tar, nicotine, and carbon monoxide are dependent upon a myriad cigarette design parameters. Although ISO generated TNCO and only nine physical features were measured for this study, many other factors, such as smoking

topography, consumer behavior factors, tobacco fillers with different tobacco cultivars, blend compositions, crop years, reconstituted or expanded tobacco, and added ingredients, were not factored into the multivariate analyses. Additional exclusions to this study included cigarette components (i.e., structural composition, features, and ingredients) that can also contribute to variations in machine-measured TNCO emissions in MSS. Thus, the PLSR models in this study cannot account for cigarettes with different filter technologies and paper types, tobacco blend compositions, ingredients, or variations in static burn rate, which impact the total puff count, even though the resulting ISO generated TNCO yields can be similar in select cases.

Marketplace survey studies involve cigarette brand styles with design properties representing the commercial cigarette market. Published market studies include typical cigarette design features on the market and are beneficial by providing a snapshot of the relative ranges of machine smoke yields.<sup>15,17–20,35</sup> This study went beyond assessing individual relationships between ISO generated TNCO yields and physical design parameters by assessing the relationships simultaneously within statistically defined clusters of cigarettes.

Initially, Pearson correlation coefficients were calculated to assess univariate relationships (see Table 2). A strong, negative relationship between ISO generated TNCO yields with filter ventilation was observed together with a strong, positive relationship among the TNCO yields. This strong, positive correlation indicates that there is a linear relationship among tar, nicotine, and CO under ISO smoking conditions and confirms what has been reported in the literature.<sup>36,37</sup> All physical parameters, with the exception of paper porosity, circumference, and draw resistance, exhibit a statistically significant correlation with ISO measured nicotine and tar emissions. Increasing cigarette paper porosity allows external air to permeate through the cigarette rod and reduces smoke from the burning coal allowing low molecular weight vapor phase constituents (i.e., CO) to diffuse through the rod into the environment and decreases measured nicotine and tar yields.<sup>37</sup> All physical parameters, with the exception of rod length, circumference, paper porosity, and pressure drop, exhibit a statistically significant correlation with ISO measured CO. Similar univariate correlations between ISO generated TNCO

and physical design parameters of commercial cigarettes are supported in the literature.<sup>22</sup>

Multivariate analyses of commercial cigarettes between mainstream machine cigarette smoke and selected smoke constituents have been conducted previously.<sup>38–40</sup> The aim of these studies was to discern the differences between cigarette brands and constituents in the particulate and vapor phases. Outcomes of many studies demonstrate correlations between tar and other HPHCs, along with a trend of increasing smoke constituent concentrations with increasing ISO measured tar yields.<sup>15,18,19</sup> Predicting HPHCs yields for an individual cigarette brand style from its ISO measured TNCO is problematic and likely inaccurate (i.e., low correlations and high prediction errors) without accounting for the smoking topographies or fully identifying all distinctive properties of a cigarette, such as variations from tobacco crop years and grades, tobacco blend composition, added ingredients, physical design parameters (e.g., filter ventilation, paper porosity, and tobacco weight), and manufacturing processes. As observed in our novel approach, when the relationship between the cigarette design parameters and the ISO generated TNCO emissions are examined simultaneously, the interactions are not always consistent between groups. Such differences could introduce unacceptable error in a prediction model fit using linear regression to models one variable (TNCO yield) at a time. In addition to differences in individual cigarettes, laboratory measurement variability can potentially decrease the ability to accurately predict HPHC yields when considering all commercial cigarettes on the market.<sup>35</sup>

A model of the relationship between the cigarette design parameters and the ISO generated TNCO emissions identified five discrete clusters illustrated in Figure 2 and listed in Table 1. Group sizes consisted of 15, 11, 1, 3, and 19 for Groups 1, 2, 3, 4, and 5, respectively. The first three PLSR components account for at least 65% of the variability in ISO TNCO yields and 80% of the variability in physical parameters in Groups 1, 2, and 5 (see Table 3). Consequently, approximately 35% of the variability in the TNCO yields is not accounted for by these nine design physical parameters. The variability calculated in this study is similar to the 75% observed in the HPHC smoke yields multivariate analysis by Piadé.<sup>40</sup> Variability in Groups 3 and 4 were not evaluated because of the limited cigarette brand styles within these clusters.

The outcome is very different from the univariate correlation results summarized in Table 2 compared to the plot (Figure 4) of the first and second PLSR loadings listed in Table 4, despite the fact that paper porosity and circumference had no statistically significant univariate correlations with the ISO generated TNCO yields. As shown by the proximity of the loadings in Figure 4, circumference is positively correlated with ISO generated CO in Group 2 and negatively correlated with nicotine and tar under ISO conditions in Group 5. Draw resistance, weakly correlated with CO based on the univariate correlation, is strongly correlated (+) with CO in Groups 1 and 2 and moderately correlated with CO in Group 5. A nearly zero univariate correlation between CO and pressure drop becomes a strong correlation (+) for Groups 1 and 2 and weak correlation (+) for Group 5 when the impact of other design parameters is taken into account.

The distribution of cigarettes within each cluster in Figure 5 demonstrates the variance in the cigarette physical features in relationship to the TNCO deliveries and range of commercial cigarettes selected in this study. Within a cluster, cigarette

brand styles were not consistently grouped by manufacturer, brand style, or length. Several cigarettes within this study can be considered outliers in Groups 1, 2, and 5, meaning their physical design features exhibit a greater influence on their TNCO deliveries. For example, filter ventilation for cigarettes 5 and 6 (Group 1) is the lowest for all cigarettes assessed; cigarette 33 (Group 2) possesses the lowest circumference, tobacco weight, pressure drop, and draw resistance; and cigarette 51, 3R4F (Group 5), is constructed with the shortest overwrap and largest circumference.

Filter ventilation, which has a strong negative univariate correlation with all three ISO generated TNCO yields, demonstrates a different trend within each group when considering the first two PLSR loadings (Figure 4). Carbon monoxide is the only yield that has a negative correlation with filter ventilation (identified as TV in Figures 4 and 5) between and within groups. In contrast, no correlation is observed among filter ventilation and tar or nicotine in Group 2; positive correlation with nicotine and no correlation with tar in Group 1; and no correlation with nicotine and negative correlation with tar in Group 5. Our results differ from previous reports where filter ventilation is the most important cigarette design feature when taken in aggregate.<sup>22,37</sup> The multivariate analysis carried out by O'Connor et al. utilized a cumulative logit model (a multiple linear regression that can lead to overfitting the data), instead of PLSR, and the origin of the cigarette samples were from U.S., U.K., Canada, and Australia where the tar yields, tobacco blends, and physical design parameters vary by country; thus, design parameters were adjusted for tar in their study, and tar was treated as a categorical outcome in Australia, instead of a continuous outcome in the other countries.<sup>22,41</sup> The O'Connor et al. study also found that tobacco weight, rod density, and overwrap were minor factors in tar yields. A direct comparison to the O'Connor et al. study results will be relative since the cigarette tobacco blends from other countries are not the same as in the U.S.

Three cigarettes clustered in Group 4 were designed with 55–67% filter ventilation, the highest in this study with the exception of cigarettes 33 and 50 in Group 2, resulting in the lowest ISO generated TNCO deliveries. Decreased ISO TNCO yields in MSS in highly ventilated cigarettes produces higher air dilution from increased ventilation holes, decreased tobacco combustion from diminished puff volume circulating through the burning zone, and lower velocity flow when drawn through the cigarette rod which results in increased CO diffusion and filtration of tar and nicotine.<sup>7,42</sup> In conjunction to higher ventilation, these cigarettes possess the lowest draw resistance, an elastic design parameter that can be manipulated by the smoker when blocking the filter ventilation holes.<sup>7,42</sup> Ventilation blocking is more prevalent in cigarettes with machine measured tar yields <2 mg/cigarette under ISO smoking conditions.<sup>7</sup> Typically, these type of cigarettes can potentially require more effort from the smoker than most smokers are willing to put forth; for example, compensatory smoking may involve blocking ventilation holes and harder, longer, more frequent puffs.<sup>14</sup> Consequently, the smoker falsely believes the “lighter” tasting (i.e., MSS diluted with a high percentage of air) and <2 mg ISO measured tar cigarette is a better choice without realizing their unconscious smoking behavior is increasing constituent exposure, particularly CO.<sup>7,43</sup> The small sample size in Group 4 precluded PLSR; however, previous studies have demonstrated inaccuracy in ISO TNCO

and smoke constituent correlations with the type of cigarettes within this group.<sup>18,37,44</sup>

As cigarettes with <4 mg of ISO measure tar yields excluded Group 4 cigarettes from Group 2, the tobacco blend separated the Group 3 cigarette from Group 1. Cigarette blends for the majority of the cigarettes in this study are blended with flue-cured, burley, and oriental tobacco leaves, and many include tobacco stems and reconstituted and expanded tobacco. The tobacco blend composition of one cigarette brand style in Group 3 is limited to primarily flue-cured tobacco, a common tobacco blend in the U.K., Australia, and Canada. Cigarettes containing only flue-cured tobacco have been observed to deliver higher polycyclic aromatic hydrocarbons yields, lower yields of tobacco-specific nitrosamines, monoaromatic volatile organic compounds, and higher cadmium levels than a corresponding blend containing primarily flue-cure and fire-cured tobaccos.<sup>40,45–47</sup>

## CONCLUSIONS

This investigation describes a novel statistical approach for evaluating the relationship between the ISO machine-measured TNCO yields and nine physical parameters using 50 U.S. cigarettes commercially available in 2011 and one reference cigarette. Our approach is unique within the tobacco literature, in that multivariate methods were used to simultaneously model the three types of mainstream ISO smoke emissions (TNCO) and the nine physical parameters. An evaluation of the influence of the nine physical design features on machine measured TNCO yields using the ISO smoking regimen first clustered the 51 cigarettes into five discrete groups based on their ISO TNCO yields. Partial least squares regression was then used to describe the correlation structure between the yields and physical parameters within each group. An important finding of this work is that the physical design features impact ISO TNCO deliveries differently within and between the groups constructed based on ISO TNCO yields. Thus, any attempt to model deliveries should incorporate this strategy and verify the application applied to appropriate brand style selections. Unknown design features and changeability in physical design parameters within the commercial cigarette marketplace and range of smoking topographies will continue to complicate the prediction of smoke constituents based on ISO generated TNCO yields.

When evaluating the 51 cigarette brand styles, the first three components (out of a maximum of nine) from the PLSR models accounted for approximately 65% of variability in the ISO generated TNCO yields within each of the clusters; some of the remaining variability can be attributed to cigarette features not included in this investigation, such as tobacco types, added ingredients, and material composition. The correlation between the predicted and observed ISO TNCO yields ranged from 0.5 to 0.9. The influence of the nine physical parameters varied between the groups and ISO generated TNCO constituent being predicted, and there was no one dominant physical parameter. Although the ISO TNCO yields are similar within each cluster, there was no discernible trends in the physical parameters of individual cigarette brand styles between clusters; variations are attributed to cigarette features. A companion paper will explore the same 51 cigarettes when TNCO yields in MSS are measured using the CI smoking conditions and investigate the variability among cigarettes in statistically defined groups for both ISO and CI. The data and analysis from this future study will help determine important

design parameters that affect TNCO yields across the spectrum of smoking topographies.

## AUTHOR INFORMATION

### Corresponding Author

\*Tel: 301-796-9373. E-mail: matthew.walters@fda.hhs.gov.

### Funding

This research was funded by the Center for Tobacco Products, U.S. Food and Drug Administration.

### Notes

The findings and conclusions in this report are those of the authors and do not necessarily represent the official views or positions of the FDA or the Department of Health and Human Services.

The authors declare no competing financial interest.

## ABBREVIATIONS

ISO, International Organization of Standardization; CI, Canadian Intense; CIR, circumference; FL, filter length; HPHC, harmful and potentially harmful constituents; MSS, mainstream smoke; OW, overwrap; PCs, principal components; PDO, pressure drop open; PDS, pressure drop shut; PLSR, partial least-squares regression; PP, paper porosity; RL, rod length; TNCO, tar, nicotine, and carbon monoxide; TV, tip ventilation; TW, tobacco weight

## REFERENCES

- (1) Federal Trade Commission (1967) *Cigarettes: Testing for Tar and Nicotine Content*, Federal Register, Vol. 32, No. 147, p 11178, Office of the Federal Register of the National Archives and Records Administration and the U.S. Government Printing Office, Washington, D.C.
- (2) Federal Trade Commission (1980) *Cigarettes and Related Matters: Carbon Monoxide, Tar and Nicotine Content of Cigarette Smoke; Description of New Machine and Methods to be Used in Testing*, Federal Register, Vol. 45, No. 134, pp 46483–46487, Office of the Federal Register of the National Archives and Records Administration and the U.S. Government Printing Office, Washington, D.C.
- (3) Commonwealth of Massachusetts (1997) 105 CMR 660.000: Cigarette and Smokeless Tobacco Products: Reports of Added Constituents and Nicotine Ratings. <http://www.mass.gov/eohhs/docs/dph/regs/105cmr660.pdf> (accessed May 2016).
- (4) Commission of the European Communities (2007) Cigarette Yield Measurement and Some Basic Steps for Laboratory Approval – Practical Guide Brussels. [http://ec.europa.eu/health/ph\\_determinants/life\\_style/Tobacco/Documents/best\\_practices\\_en.pdf](http://ec.europa.eu/health/ph_determinants/life_style/Tobacco/Documents/best_practices_en.pdf) (accessed May 2016).
- (5) U.S. Department of Health and Human Services (2010) *How Tobacco Smoke Causes Disease: The Biology and Behavioral Basis for Smoking-Attributable Disease: A Report of the Surgeon General*, pp 13–22, Centers for Disease Control and Prevention, National Center for Chronic Disease Prevention and Health Promotion, Office on Smoking and Health (US), Washington, D.C. <http://www.ncbi.nlm.nih.gov/books/NBK53013/> (accessed May 2016).
- (6) Hoffmann, D., and Hoffmann, I. (2001) The Changing Cigarette: Chemical Studies and Bioassays, in *Tobacco Monograph 13: Risks Associated with Smoking Cigarettes with Low Machine-Measured Yield of Tar and Nicotine*, National Institutes of Health, National Cancer Institute, Division of Cancer Control and Population Sciences, Bethesda, MD. [http://cancercontrol.cancer.gov/brp/tcrb/monographs/13/m13\\_5.pdf](http://cancercontrol.cancer.gov/brp/tcrb/monographs/13/m13_5.pdf) (accessed May 2016).
- (7) Kozlowski, L. T., and O'Connor, R. J. (2002) Cigarette Filter Ventilation is a Defective Design because of Misleading Taste, Bigger Puffs, and Blocked Vents. *Tob. Control* 11 (Suppl 1), I40–I50.
- (8) Food and Drug Administration (2009) Family Smoking Prevention and Tobacco Control Act, In Public Law 111–31,

Rockville, MD. <http://www.gpo.gov/fdsys/pkg/PLAW-111publ31/pdf/PLAW-111publ31.pdf> (accessed May 2016).

(9) Food and Drug Administration (2012) Reporting Harmful and Potentially Harmful Constituents in Tobacco Products and Tobacco Smoke under Section 904(a)(3) of the Federal Food, Drug, and Cosmetic Act, Rockville, MD. <http://www.fda.gov/TobaccoProducts/GuidanceComplianceRegulatoryInformation/ucm297752.htm> (accessed May 2016).

(10) Brazilian Health Surveillance Agency (ANVISA) (2007) Resolution - RDC No. 90, Brasilia, Brasil. <http://www.tobaccocontrol.org/files/live/Brazil/Brazil%20-%20RDC%20No.%2090.pdf> (accessed May 2016).

(11) Australian Government Department of Health (2001) Cigarette Emissions Data, Emissions Data for Selected Australian Cigarette Brand Variants, Canberra, Australia. <http://www.health.gov.au/internet/main/publishing.nsf/Content/tobacco-emis> (accessed May 2016).

(12) Taiwan Department of Health (2015) Regulations Governing Reporting of Tobacco Product Information under the Tobacco Hazards Prevention & Control Act (THPCA), Health Promotion Administration, Ministry of Health and Welfare, Taipei City, Taiwan. <http://www.hpa.gov.tw/English/ClassShow.aspx?No=200907130001> (accessed May 2016).

(13) Health Canada (2005) Tobacco Reporting Regulations (TTR) (2000) and Its Amendments (2005), Health Canada. <http://www.hc-sc.gc.ca/hc-ps/tobac-tabac/legislation/reg/indust/index-eng.php> (accessed May 2016).

(14) Scherer, G. (1999) Smoking Behaviour and Compensation: a Review of the Literature. *Psychopharmacology (Heidelberg, Ger.)* 145, 1–20.

(15) Counts, M. E., Hsu, F. S., Laffoon, S. W., Dwyer, R. W., and Cox, R. H. (2004) Mainstream Smoke Constituent Yields and Predicting Relationships from a Worldwide Market Sample of Cigarette Brands: ISO Smoking Conditions. *Regul. Toxicol. Pharmacol.* 39, 111–134.

(16) Calafat, A. M., Polzin, G. M., Saylor, J., Richter, P., Ashley, D. L., and Watson, C. H. (2004) Determination of Tar, Nicotine, and Carbon Monoxide Yields in the Mainstream Smoke of Selected International Cigarettes. *Tob. Control* 13, 45–51.

(17) Borgerding, M. F., Bodnar, J. A., and Wingate, D. E. (2000) The 1999 Massachusetts Benchmark Study – Final Report: Presented to the Massachusetts Department of Public Health. <http://legacy.library.ucs.edu/tid/yek21c00> (accessed May 2016).

(18) Chepiga, T. A., Morton, M. J., Murphy, P. A., Avalos, J. T., Bombick, B. R., Doolittle, D. J., Borgerding, M. F., and Swauger, J. E. (2000) A Comparison of the Mainstream Smoke Chemistry and Mutagenicity of a Representative Sample of the US Cigarette Market with Two Kentucky Reference Cigarettes (K1R4F and K1R5F). *Food Chem. Toxicol.* 38, 949–962.

(19) Swauger, J. E., Steichen, T. J., Murphy, P. A., and Kinsler, S. (2002) An Analysis of the Mainstream Smoke Chemistry of Samples of the U.S. Cigarette Market Acquired between 1995 and 2000. *Regul. Toxicol. Pharmacol.* 35, 142–156.

(20) Bodnar, J. A., Morgan, W. T., Murphy, P. A., and Ogden, M. W. (2012) Mainstream Smoke Chemistry Analysis of Samples from the 2009 US Cigarette Market. *Regul. Toxicol. Pharmacol.* 64, 35–42.

(21) Adam, T., McAughey, J., Mocker, C., McGrath, C., and Zimmermann, R. (2010) Influence of Filter Ventilation on the Chemical Composition of Cigarette Mainstream Smoke. *Anal. Chim. Acta* 657, 36–44.

(22) O'Connor, R. J., Hammond, D., McNeill, A., King, B., Kozlowski, L. T., Giovino, G. A., and Cummings, K. M. (2008) How do Different Cigarette Design Features Influence the Standard Tar Yields of Popular Cigarette Brands Sold in Different Countries? *Tob. Control* 17 (Suppl 1), i1–i5.

(23) Connolly, G. N., Alpert, H. R., Wayne, G. F., and Koh, H. (2007) Trends in Nicotine Yield in Smoke and Its Relationship with Design Characteristics among Popular US Cigarette Brands, 1997–2005. *Tob. Control* 16, e5.

(24) Maxwell, J. C. (2008) *The Maxwell Report: Year End & Fourth Quarter 2007 Sales Estimates for the Cigarette Industry*, John C. Maxwell, Jr., Richmond, VA.

(25) International Organization for Standardization (1999) *ISO 3402:1999 Tobacco and Tobacco Products-Atmosphere for Conditioning and Testing*, International Organization for Standardization, London, U.K.

(26) International Organization for Standardization (2000) *ISO 3308:2000 Routine Analytical Cigarette-Smoking Machine-Definitions and Standard Conditions*, International Organization for Standardization, London, U.K.

(27) International Organization for Standardization (2009) *ISO 8454:2009 Cigarettes-Determination of Carbon Monoxide in the Vapour Phase of Cigarette Smoke-NDIR Method*, International Organization for Standardization, London, U.K.

(28) International Organization for Standardization (2000) *ISO 10315:2000 Cigarettes-Determination of Nicotine in Smoke Condensates-Gas Chromatographic Method*, International Organization for Standardization, London, U.K.

(29) International Organization for Standardization (1999) *ISO 10362-1:1999 Cigarettes-Determination of Water in Smoke Condensates-Part 1: Gas Chromatographic Method*, International Organization for Standardization, London, U.K.

(30) International Organization for Standardization (2008) *ISO 4387:2008 Cigarettes-Determination of Total and Nicotine-Free Dry Particulate Matter Using a Routine Analytical Smoking Machine*, International Organization for Standardization, London, U.K.

(31) International Organization for Standardization (2013) *ISO 6565:2011 Tobacco and Tobacco Products: Draw Resistance of Cigarettes and Pressure Drop of Filter Rods: Standard Conditions and Measurement*, International Organization for Standardization, London, U.K.

(32) International Organization for Standardization (2013) *ISO 2971:2013 Cigarettes and Filter Rods: Determination of Nominal Diameter: Method Using a Non-Contact Optical Measuring Apparatus*, International Organization for Standardization, London, U.K.

(33) CORESTA. (1994) CORESTA Recommended Method No 40: Determination of Air Permeability of Materials Used as Cigarette Papers, Filter Plug Wrap and Filter Joining Paper Including Materials Having an Oriented Permeability Zone, Paris, France. <https://www.coresta.org/determination-air-permeability-materials-used-cigarette-papers-filter-plug-wrap-and-filter-joining> (accessed May 2016).

(34) International Organization for Standardization (2009) *ISO 2965:2009 Materials Used as Cigarette Papers, Filter Plug Wrap and Filter Joining Paper, Including Materials Having a Discrete or Oriented Permeable Zone and Materials with Bands of Differing Permeability: Determination of Air Permeability*, International Organization for Standardization, London, U.K.

(35) Purkis, S. W., Meger, M., and Wuttke, R. (2012) A Review of Current Smoke Constituent Measurement Activities and Aspects of Yield Variability. *Regul. Toxicol. Pharmacol.* 62, 202–213.

(36) King, B., Borland, R., Abdul-Salaam, S., Polzin, G., Ashley, D., Watson, C., and O'Connor, R. J. (2010) Divergence between Strength Indicators in Packaging and Cigarette Engineering: a Case Study of Marlboro Varieties in Australia and the USA. *Tob. Control* 19, 398–402.

(37) Stephens, W. E. (2007) Dependence of Tar, Nicotine and Carbon Monoxide Yields on Physical Parameters: Implications for Exposure, Emissions Control and Monitoring. *Tob. Control* 16, 170–176.

(38) Perez-Bernal, J. L., Amigo, J. M., Fernandez-Torres, R., Bello, M. A., and Callejon-Mochon, M. (2011) Trace-Metal Distribution of Cigarette Ashes as Marker of Tobacco Brands. *Forensic Sci. Int.* 204, 119–125.

(39) Marcilla, A., Martinez, I., Berenguer, D., Gomez-Siurana, A., and Beltran, M. I. (2012) Comparative Study of the Main Characteristics and Composition of the Mainstream Smoke of Ten Cigarette Brands Sold in Spain. *Food Chem. Toxicol.* 50, 1317–1333.

(40) Piadé, J.-J., Wajrock, S., Jaccard, G., and Janeke, G. (2013) Formation of Mainstream Cigarette Smoke Constituents Prioritized by

the World Health Organization–Yield Patterns Observed in Market Surveys, Clustering and Inverse Correlations. *Food Chem. Toxicol.* 55, 329–347.

(41) Wold, S., Sjöström, M., and Eriksson, L. (2001) PLS-Regression: a Basic Tool of Chemometrics. *Chemom. Intell. Lab. Syst.* 58, 109–130.

(42) Baker, R. R. (2006) Smoke Generation Inside a Burning Cigarette: Modifying Combustion to Develop Cigarettes that May Be Less Hazardous to Health. *Prog. Energy Combust. Sci.* 32, 373–385.

(43) Sweeney, C. T., Kozlowski, L. T., and Parsa, P. (1999) Effect of Filter Vent Blocking on Carbon Monoxide Exposure from Selected Lower Tar Cigarette Brands. *Pharmacol., Biochem. Behav.* 63, 167–173.

(44) Sweeney, C. T., and Kozlowski, L. T. (1998) Blocking Filter Vents Increases Carbon Monoxide Levels from Ultralight, but not Light Cigarettes. *Pharmacol., Biochem. Behav.* 59, 767–773.

(45) Ding, Y. S., Zhang, L., Jain, R. B., Wang, R. Y., Ashley, D. L., and Watson, C. H. (2008) Levels of Tobacco-Specific Nitrosamines and Polycyclic Aromatic Hydrocarbons in Mainstream Smoke from Different Tobacco Varieties. *Cancer Epidemiol., Biomarkers Prev.* 17, 3366–3371.

(46) Pappas, R. S., Fresquez, M. R., Martone, N., and Watson, C. H. (2014) Toxic Metal Concentrations in Mainstream Smoke from Cigarettes Available in the USA. *J. Anal. Toxicol.* 38, 204–211.

(47) Pazo, D. Y., Moliere, F., Sampson, M. M., Reese, C. M., Agnew-Heard, K. A., Walters, M. J., Holman, M. R., Blount, B. C., Watson, C., and Chambers, D. M. (2016) Mainstream Smoke Levels of Volatile Organic Compounds in 50 US Domestic Cigarette Brands Smoked with the ISO and Canadian Intense Protocols. *Nicotine Tob. Res.*