

Original Investigation

Interlaboratory comparability of serum cotinine measurements at smoker and nonsmoker concentration levels: A round-robin study

John T. Bernert, Peyton Jacob III, David B. Holiday, Neal L. Benowitz, Connie S. Sosnoff, Mira V. Doig, Colin Feyerabend, Kenneth M. Aldous, Mehran Sharifi, Mark D. Kellogg, & Loralie J. Langman

Abstract

Introduction: Cotinine, the primary proximate metabolite of nicotine, is commonly measured as an index of exposure to tobacco in both active users of tobacco and nonsmokers with possible exposure to secondhand smoke (SHS). A number of laboratories have implemented analyses for measuring serum cotinine in recent years, but there have been few interlaboratory comparisons of the results. Among nonsmokers exposed to SHS, the concentration of cotinine in blood can be quite low, and extensive variability in these measurements has been reported in the past.

Methods: In this study, a group of seven laboratories, all experienced in serum cotinine analysis, measured eight coded serum pools with concentrations ranging from background levels of about 0.05 ng/ml to relatively high concentrations in the active smokers range. All laboratories used either gas–liquid chromatography with nitrogen–phosphorus detection or liquid chromatography with mass spectrometric detection.

Results: All seven laboratories reliably measured the cotinine concentrations in samples that were within the range of their methods. In each case, the results for the pools were correctly ranked in

order, and no significant interlaboratory bias was observed at the 5% level of significance for results from any of the pools.

Discussion: We conclude that present methods of chromatographic analysis of serum cotinine, as used by these experienced laboratories, are capable of providing accurate and precise results in both the smoker and the nonsmoker concentration range.

Introduction

The use of tobacco remains an important public health problem throughout the world. Although the number of active smokers has declined in recent years, it is currently estimated that approximately 21% of the U.S. adult population continue to smoke cigarettes (Centers for Disease Control and Prevention, 2007); a similar prevalence of smoking exists in Canada (Gilmore, 2002) and in Europe (Strong et al., 2008). Furthermore, the use of cigarettes and other forms of tobacco is increasing worldwide, partly in concert with population growth, with a notable increase among females in developing countries (Mackay, Eriksen, & Shafey, 2006). As a consequence, health

John T. Bernert, Ph.D., *Division of Laboratory Science, National Center for Environmental Health, Centers for Disease Control and Prevention, Atlanta, GA*

Peyton Jacob, III, Ph.D., *Department of Medicine, University of California at San Francisco*

David B. Holiday, Ph.D., *RTI International, Atlanta Regional Office, Atlanta, GA*

Neal L. Benowitz, M.D., *Department of Medicine, University of California at San Francisco*

Connie S. Sosnoff, M.S., *Division of Laboratory Science, National Center for Environmental Health, Centers for Disease Control and Prevention, Atlanta, GA*

Mira V. Doig, Ph.D., *ABS Laboratories Ltd, Welwyn Garden City, Herts, UK*

Colin Feyerabend, Ph.D., *ABS Laboratories Ltd, Welwyn Garden City, Herts, UK*

Kenneth M. Aldous, Ph.D., *Biggs Laboratory, Wadsworth Center, New York State Department of Health, Empire State Plaza, Albany*

Mehran Sharifi, Ph.D., *LabStat International ULC, Kitchener, Ontario, Canada*

Mark D. Kellogg, Ph.D., *Department of Laboratory Medicine, Children's Hospital Boston, Boston, MA*

Loralie J. Langman, Ph.D., *Department of Laboratory Medicine and Pathology, Mayo Clinic, Rochester, MN*

Corresponding Author:

John T. Bernert, *Centers for Disease Control and Prevention, 4770 Buford Highway, N.E., Mailstop F-47, Atlanta, GA 30341, USA. Telephone: 770-488-7911; Fax: 770-488-0181; E-mail: jtb2@cdc.gov*

doi: 10.1093/ntr/ntp161

Advance Access published on November 23, 2009

Received May 6, 2009; accepted September 22, 2009

Published by Oxford University Press on behalf of the Society for Research on Nicotine and Tobacco 2009.

effects associated with the use of tobacco are on the increase in developing regions of the world. The exposure of nonsmokers to secondhand smoke (SHS) also remains a significant health problem. Despite recent substantial progress in reducing the exposure of nonsmokers to SHS (Borland, Mullins, Trotter, & White, 1999; Jarvis et al., 2000; Pirkle, Bernert, Caudill, Sosnoff, & Pechacek, 2006), such exposures continue to present a significant public health risk (Department of Health and Human Services, 2006). Assessing the exposure of nonsmokers to SHS can be especially difficult, since people may differ in their awareness of and sensitivity to such exposures in their daily lives. Currently, the most suitable marker for identifying tobacco use or for SHS exposure assessments is the measurement of cotinine, the primary proximate metabolite of nicotine, in blood (Benowitz, 1996).

Interest in monitoring SHS exposure has led to the development of new analytical methods for cotinine measurements and the involvement of several laboratories in these assays, but there have been few interlaboratory comparisons of cotinine analyses. Biber et al. (1987) evaluated nicotine and cotinine measurements in both serum and urine in a group of 11 laboratories from six countries using either radioimmunoassay (RIA) or gas chromatographic (GC) procedures. They found that all the laboratories in their study could reliably distinguish between samples from regular smokers and samples from nonsmokers. In addition, all the laboratories provided consistent ranking of results for serum cotinine. However, there were often substantial differences in the concentrations measured by the participating laboratories in that study. These authors concluded that whereas classification of smoker versus nonsmoker by cotinine analysis was generally reliable, the individual values measured by different laboratories could vary considerably. There were also substantial differences noted in that study between RIA and GC results in urine samples, although not in serum, probably reflecting the influence of possible cross-reaction in the RIA with cotinine glucuronides and 3-hydroxycotinine in the urine measurements.

More recently, a common set of urine samples was evaluated for their cotinine concentrations by a group of six European laboratories using high-performance liquid chromatography-ultraviolet (HPLC-UV) and immunoassay methods; generally, good agreement was noted within the group (Jacob et al., 2005). However, in that study, all participating laboratories were required to use the same standardized procedure for the HPLC-UV assays. As in the study by Biber et al. (1987), some differences were noted between liquid chromatography (LC) and immunoassay methods, presumably again involving the cross-reactivity of other analytes in the urine samples during the immunoassays, especially 3-hydroxycotinine.

Current analyses of cotinine in samples from nonsmokers may place increasingly greater stress on the assays because the concentration levels are often quite low. Nevertheless, maintaining good accuracy and precision in these assays is important for facilitating comparisons among studies. In some cases, even relatively small mean concentration differences between groups of nonsmokers may be of interest. Fortunately, there have been improvements in the available analytical methodology in recent years. Current assays using capillary GC with nitrogen-phosphorus detectors (GC/NPD) or, in some cases, mass spectrometric (GC/MS) detection or LC coupled with tandem mass spectrometry (LC/MS/MS) may be expected to provide both for more sensitive and more precise measurements than were avail-

able in the past. To evaluate the current state of serum cotinine measurements in smokers and nonsmokers with a particular emphasis on the latter, we have undertaken a study of the analysis of a common set of serum samples by seven laboratories experienced in these assays. Our objectives in this work were to assess the current state of cotinine analyses using modern analytical techniques and to provide consensus values for a set of serum samples that might then be used in support of further method validation activities in the future.

Materials and methods

Preparation of cotinine perchlorate

Cotinine perchlorate was used in some cases to fortify selected human serum pools to achieve known concentrations of cotinine. The perchlorate salt was prepared to further purify the analytical standard and to provide a material that was less hygroscopic and more easily weighed than free-base cotinine. Cotinine perchlorate was prepared essentially as described by Hariharan, VanNoord, and Greden (1988). The original cotinine stock was Sigma (St. Louis, MO) C-5923, lot number 082K4020, with a stated purity of approximately 98%. The perchloric acid used was Aldrich (Milwaukee, WI) 31,142-1, listed as 99.999% pure (based on metals content). All other reagents were of the highest purity available. The cotinine free base was dissolved in isopropanol and perchloric acid was added. The initial white precipitate was collected and dried under vacuum and then recrystallized from methanol. After recovery and drying, the crystals were redissolved in methanol, treated with a small amount of decolorizing carbon (Darco G-60, Marshall, TX), and then recrystallized two more times. The final material was dried under vacuum and stored in a desiccator at room temperature.

Aliquots of the final product were sent to Galbraith Laboratories, Knoxville, TN, for melting point measurements and elemental analysis. The results of those assays, calculated as an anhydrous monoperchlorate cotinine salt, were as follows: % carbon found 43.58, calculated as 43.41; % hydrogen found 4.71, calculated as 4.73; % nitrogen found 10.01, calculated as 10.12; and % chlorine found 12.64, calculated as 12.81. The melting point was 218.5–219.3°C. Additional aliquots were analyzed by proton nuclear magnetic resonance in deuterium oxide in the Analytical Chemistry Department of the Georgia Institute of Technology, Atlanta. The apparent estimated purity of the product from those measurements was >99%. Solutions of the compound were prepared, basified, and extracted and then analyzed by LC atmospheric pressure ionization (API) MS/MS and by capillary GC MS on both single quadrupole and magnetic sector (high resolution) instruments. A single peak for cotinine was seen in each case, with no additional contaminants noted. Aliquots of this final product in water were added to a serum pool base obtained from nonsmokers with little or no known SHS exposure and thoroughly mixed to produce samples at known concentrations.

Serum pools

Pools were prepared in Atlanta, Georgia, from human serum collected from both smokers and nonsmokers. Eight human serum pools were evaluated in this study, as summarized in Table 1. The base pool E was prepared from serum obtained from a group of nonsmokers who did not live or work with smokers and who

Table 1. Serum pools examined

Pool ID	Fortified	Preparation	Target concentration
A	Yes	0.461 ng/ml ^a	0.511 ng/ml ^b
B	Yes	1.92 ng/ml ^a	1.97 ng/ml
C	Yes	46.1 ng/ml ^a	46.2 ng/ml
D	Yes	192 ng/ml ^a	192 ng/ml
E	No	Base pool	Consensus ^c
F	No	Nonsmoker pool	Consensus
G	No	Blended pool	Consensus
H	No	Smoker pool	Consensus

Note. ^aAn aliquot of the cotinine perchlorate stock was added to fortified pools at a concentration of the salt calculated to produce this final predicted concentration of cotinine as the free base.

^bThe target concentration was calculated by adding the consensus mean value for the base pool (Pool E; 0.05 ng/ml) as measured by Laboratories 2, 6, and 7 to the initial fortified concentration.

^cPool E was the base pool used for preparation of Pools A–D. It was obtained from a selected group of nonsmokers with no known secondhand smoke exposure.

had no known exposure to SHS. The results from measurements of this base pool, analyzed as unknowns by the three participating laboratories capable of measuring the pool, were used to provide a consensus baseline concentration that was included in the target concentration for the fortified pools. The four fortified pools (A–D) were prepared to provide target concentrations in both the smoker and the nonsmoker range from approximately 0.5 to 190 ng/ml as free-base cotinine (Table 1).

Three additional unmodified human serum pools (F–H) were included in the study: two with relatively low cotinine concentrations, consistent with SHS-exposed nonsmokers, and one in the active smoker concentration range. After extensive mixing, each of the eight pools was aliquoted into cryovials, which were labeled with a set of randomly generated four-digit numbers. A set of aliquots was then selected at random from each pool and analyzed in a single run to confirm homogeneity. Laboratories were blinded to the identity of the samples, although the individual vial ID numbers belonging to a particular pool were identified in advance. Prior studies suggest that distinguishing between smoker- and nonsmoker-level samples is relatively trivial when using current chromatographic assays. Thus, to reduce the burden on the participating laboratories by eliminating preclassification time, the smoker-level pools (Pools C, D, and H) were also identified. An aqueous stock of cotinine perchlorate at an expected concentration by weight of 1.505 µg of cotinine perchlorate/ml (estimated as 960.3 ng/ml of cotinine as the free base) was also provided to each laboratory. The expected concentration was indicated on that sample, and each laboratory was encouraged to use this stock solution as it saw fit.

Analytical procedures

A summary of the analytical methods used by each of the laboratories is given in Table 2. All seven participating laboratories used chromatographic methods for these serum cotinine assays. Five laboratories used some form of LC API MS/MS, one laboratory used GC/NPD (GC with thermionic detection), and one

Table 2. Analytical methods

Laboratory	Method	Internal standard	Detection limit (ng/ml)
1-A	Capillary GLC/NPD	5-Methylcotinine	0.1
1-B	HPLC API MS/MS	Cotinine-D ₃	<0.1
2	HPLC APCI MS/MS	Cotinine-D ₃	<0.1
3	HPLC ESI MS/MS	Cotinine-D ₃	0.5
4	Capillary GLC/NPD	5-Methylcotinine	3.4 ^a
5	HPLC API MS/MS	Cotinine-D ₃	2
6	HPLC API MS/MS	Cotinine-D ₃	<0.1
7	HPLC APCI MS/MS	Cotinine-D ₉	<0.1

Note. GLC = gas-liquid chromatography; NPD = nitrogen-phosphorous specific detector (thermionic detection); HPLC = high-performance liquid chromatography; API = atmospheric pressure ionization (generic); APCI = atmospheric pressure chemical ionization; ESI = electrospray ionization; and MS/MS = tandem mass spectrometry.

^aThe value given is the reported limit of quantitation; the limit of detection was 1.0.

laboratory (Laboratory 1) reported results using both an LC/MS/MS and a GC/NPD method. Both GC/NPD methods in this study used 5-methylcotinine as the internal standard, and all LC/MS/MS methods used trideuterated cotinine as the internal standard, except for Laboratory 7, which used cotinine-D₉. No immunoassays were used for these analyses. The reported limit of detection (LOD) for the various methods ranged from <0.1 to 2 ng/ml; one laboratory censored results that were above its LOD (1 ng/ml) but less than its limit of quantification (3.4 ng/ml), and we used 3.4 ng/ml as the effective quantification cutoff for that laboratory. Each laboratory analyzed the samples as unknowns, according to its standard procedure. The participating laboratories were asked to assay duplicate aliquots from each pool in runs conducted on three separate days. No time deadlines were established for these analyses, and the laboratories were free to mix these samples in with their ongoing studies. Coded samples were shipped from the coordinating laboratory over dry ice in late April 2008, and all results were obtained by November 2008.

Statistical analysis

We followed the definitions and procedures outlined in International Standardization Organization (ISO) documents (ISO 5725-1, ISO 5725-2, and ISO 5725-4) except as indicated below. The first and last of these documents define *trueness* to describe the closeness of agreement between the arithmetic mean of a large number of test results and the true or accepted reference values (our targets for the pools in Table 1). *Precision* refers to the closeness of agreement between test results. An estimate of *bias* from an assessing laboratory is the difference between the mean of the pool target concentration and its accepted target reference value. To assess the significance of the bias estimates, we constructed 95% CIs around this difference, using equations from section 4.7.2 in ISO 5725-4:1994(E). No significant bias was declared if the interval covered the value 0.

Other concepts in the context of laboratory testing include *repeatability*, which characterizes variability among replicates obtained in the same laboratory on the same material (target concentration), and the *between-laboratory variance*, which

Table 3. All serum cotinine data (ng/ml)

Pool	Sample	Laboratory							
		1-A	1-B	2	3	4	5	6	7
A	1	0.60	0.424	0.490	–	–	–	0.518	0.535
	2	0.64	0.414	0.494	–	–	–	0.519	0.541
	3	0.54	0.424	0.496	–	–	–	0.430	0.545
	4	0.40	0.414	0.479	–	–	–	0.526	0.563
	5	0.58	0.414	0.499	–	–	–	0.444	0.559
	6	0.62	0.424	0.492	–	–	–	0.523	0.558
B	1	1.90	1.874	1.844	1.420	–	–	1.969	2.109
	2	2.19	1.854	1.863	1.490	–	–	1.929	2.035
	3	2.23	1.864	1.920	1.260	–	–	1.949	2.089
	4	2.01	1.854	1.818	1.330	–	–	1.989	2.156
	5	2.31	1.954	1.928	1.440	–	–	1.939	2.112
	6	2.19	1.934	1.868	1.140	–	–	1.829	2.160
C	1	48.39	45.57	44.68	38.10	43.5	41.87	44.38	48.16
	2	48.56	46.26	45.25	39.60	45.0	42.87	45.68	52.03
	3	48.79	47.02	44.42	42.40	37.8	42.85	47.08	43.15
	4	48.57	45.45	45.63	42.60	42.0	41.85	47.28	44.34
	5	49.16	46.28	46.86	41.30	36.4	41.00	45.78	45.17
	6	48.49	46.82	46.36	41.00	40.9	43.00	45.48	46.40
D	1	200.1	205.1	180.8	172.0	180.9	175.9	189.0	173.9
	2	200.9	216.9	185.0	163.0	177.7	178.9	185.0	215.0
	3	201.2	210.2	175.9	168.0	190.7	172.9	186.0	200.4
	4	196.2	209.8	187.9	177.0	170.6	176.9	179.0	184.7
	5	196.1	217.0	190.7	160.0	174.0	172.0	188.0	204.1
	6	198.0	218.4	190.5	174.0	185.9	182.0	190.0	192.0
E	1	–	–	0.047	–	–	–	0.047	0.050
	2	–	–	0.051	–	–	–	0.046	0.058
	3	–	–	0.051	–	–	–	0.056	0.050
	4	–	–	0.050	–	–	–	0.053	0.048
	5	–	–	0.050	–	–	–	0.052	0.055
	6	–	–	0.054	–	–	–	0.055	0.054
F	1	–	–	0.112	–	–	–	0.114	0.123
	2	–	–	0.109	–	–	–	0.113	0.125
	3	–	–	0.110	–	–	–	0.106	0.117
	4	–	–	0.104	–	–	–	0.102	0.119
	5	–	–	0.112	–	–	–	0.122	0.121
	6	–	–	0.116	–	–	–	0.100	0.123
G	1	1.20	1.254	1.246	0.904	–	–	1.299	1.384
	2	1.33	1.214	1.202	1.010	–	–	1.259	1.337
	3	1.36	1.224	1.240	0.866	–	–	1.089	1.352
	4	1.26	1.254	1.258	0.863	–	–	1.309	1.342
	5	1.34	1.224	1.242	0.883	–	–	1.029	1.402
	6	1.27	1.244	1.193	0.680	–	–	1.229	1.344
H	1	200.4	205.3	180.1	184.0	169.8	164.9	176.0	189.3
	2	201.5	204.6	179.1	184.0	175.8	168.9	176.0	195.3
	3	200.2	202.5	181.4	174.0	154.1	180.9	184.0	183.6
	4	200.8	203.8	180.0	174.0	159.2	182.9	185.0	182.0
	5	200.9	198.8	183.7	169.0	160.1	161.0	177.0	191.9
	6	200.3	201.2	176.0	166.0	158.1	166.0	180.0	179.8

refers to variability among laboratories using the same method. We used the approach outlined in ISO 5725-2:1994(E) to assess these characteristics (equations 7.4.5.1 and 7.4.5.2, respectively). The sum of these two quantities is called the *reproducibility variance* (equation 7.4.5.5). We also examined the functional relationship between precision values and the mean level for each

of the target level samples, using ISO 5725-2:1994(E) following section 7.5. It is not unusual for the repeatability (S_r) and reproducibility (S_R) variances to follow a linear relationship with the mean values, often through the origin. We tested whether an intercept other than the origin was needed, and we excluded the intercept in the model if it was deemed nonsignificant.

Table 4. Summary results by laboratory and pool (ng/ml)

Laboratory, M (SD) range	Pool							
	1-A	1-B	2	3	4	5	6	7
A	0.56 (0.087) 0.40-0.64	0.42 (0.007) 0.41-0.43	0.49 (0.007) 0.48-0.50	-	-	-	0.49 (0.044) 0.43-0.53	0.55 (0.012) 0.53-0.56
B	2.14 (0.154) 1.90-2.31	1.89 (0.045) 1.85-1.96	1.87 (0.043) 1.82-1.93	1.35 (0.130) 1.14-1.49	-	-	1.93 (0.056) 1.83-1.99	2.11 (0.046) 2.03-2.16
C	48.66 (0.278) 48.4-49.2	46.24 (0.635) 45.5-47.0	45.5 (0.949) 44.4-46.9	40.83 (1.721) 38.1-42.6	40.93 (3.306) 36.4-45.0	42.24 (0.797) 41.0-43.0	45.95 (1.080) 44.4-47.3	46.54 (3.193) 43.2-52.0
D	198.8 (2.306) 196.1-201.2	212.9 (5.311) 205.1-218.4	185.1 (5.87) 175.9-190.7	169.0 (6.573) 160.0-177.0	180.0 (7.481) 170.6-190.7	176.4 (3.739) 172.0-182.0	186.2 (3.971) 179.0-190.0	195.0 (14.63) 173.9-215.0
E	-	-	0.05 (0.002) 0.05-0.05	-	-	-	0.05 (0.004) 0.05-0.06	0.05 (0.004) 0.05-0.06
F	-	-	0.11 (0.004) 0.10-0.12	-	-	-	0.11 (0.008) 0.10-0.12	0.12 (0.003) 0.12-0.12
G	1.29 (0.061) 1.20-1.36	1.24 (0.016) 1.22-1.25	1.23 (0.026) 1.19-1.26	0.87 (0.107) 0.68-1.01	-	-	1.20 (0.116) 1.03-1.31	1.36 (0.027) 1.34-1.40
H	200.7 (0.479) 200.2-201.5	202.7 (2.433) 198.8-205.3	180.1 (2.55) 176.0-183.7	175.2 (7.494) 166.0-184.0	162.9 (8.198) 154.1-175.8	170.7 (8.991) 161.0-182.9	179.7 (4.033) 176.0-185.0	187.0 (6.141) 179.8-195.4

Table 5. Cotinine level in serum: Estimation of repeatability and reproducibility standard deviations and bias of the measurement method

Pool	Laboratory							
	A	B	C	D	E	F	G	H
n	6	6	6	6	6	6	6	6
p	5	6	8	8	3	3	6	8
S_r	0.0442	0.0912	1.8463	7.1678	0.0035	0.0057	0.0710	5.8183
S_k	0.0697	0.2971	3.3530	15.3823	0.0033	0.0082	0.1833	14.8721
γ	1.5787	3.2577	1.8161	2.1460	0.9538	1.4462	2.5813	2.5561
A	0.7151	0.7681	0.5991	0.6271	0.3278	0.8777	0.7485	0.6473
A_{S_R}	0.0498	0.2282	2.0087	9.6469	0.0011	0.0072	0.1372	9.6262
\bar{y}	0.5040	1.8820	44.6150	187.9150	0.0510	0.1140	1.1980	182.3560
μ	0.5110	1.9700	46.2000	192.0000	N/A	N/A	N/A	N/A
$\hat{\delta}$	-0.0070	-0.0880	-1.5850	-4.0850	N/A	N/A	N/A	N/A
$\hat{\delta} - A_{S_R}$	-0.0568	-0.3162	-3.5937	-13.7319	N/A	N/A	N/A	N/A
$\hat{\delta} + A_{S_R}$	0.0428	0.1402	0.4237	5.5619	N/A	N/A	N/A	N/A

Note: All quantities calculated from pages 3-5 of ISO 5725-4:1994(E) with formula references below: n = number of test results within pool and laboratory; p = number of laboratories; S_r = square root of the repeatability variance (formula 8); S_k = square root of the reproducibility variance (formula 12); $\gamma = S_r/S_k$ (under formula 18); A = a function of p , n , and γ (formula 6); $A_{S_R} = A \times S_R$; \bar{y} = mean of the p laboratory means (formula 13); μ = target values by pools; $\hat{\delta}$ = estimate of the bias ($\bar{y} - \mu$; formula 15); $\hat{\delta} - A_{S_R}$ = lower confidence interval (formula 18); $\hat{\delta} + A_{S_R}$ = upper confidence interval (formula 18); N/A = not available.

In examining precision according to ISO 5725-2:1994(E), we first obtained an overview of the data, using Mandel's h and k charts before more detailed analyses. The statistic h is calculated for each participant laboratory and pool combination by a standardization of the mean of the replicate measurements for each laboratory/pool. The average of the replicates for all laboratories reporting on the pool is subtracted and the difference is divided by the standard deviation, again calculated over all participants for the given pool. Note that h is either positive or negative when the laboratory obtains, on the average, a value higher (or lower) than the average of all laboratories for the given pool. A plot was made for all the h values for all pools, arranged by laboratories.

The statistic k measures the relative variability between the replicate measurements of any laboratory for a given pool. It is the ratio of the standard deviation among the replicates for a given laboratory and serum pool to a statistically pooled standard deviation, which is calculated as the square root of the average of all laboratories reporting on each material pool in turn. A value of $k > 1$ indicates a laboratory whose replicates vary more than the average. A value of $k < 1$ indicates that the laboratory has lower-than-average interreplicate variability. These values are plotted in a similar way to the h values. Critical values can be obtained to quantify significant departures from average behavior at a chosen level of significance. Because this was designed as a descriptive study with a small number of laboratories per method, we did not stratify by different laboratory methods in the analysis.

We also performed Cochran's test for variance heterogeneity, but we will not report these statistics because most of the variance can be attributed to different methods rather than to experimental variability. Also, because of the nature of the experimental design, no outlier investigation was conducted and all observations were kept in the final analysis. Therefore, we have not reported any outcomes for Grubbs' test, which looks for outlier laboratory averages. Outlier exclusions for any reason are not generally recommended (Mandel, 1998).

Results

All measured values (blank adjusted as necessary) reported by the laboratories participating in this study are given in Table 3. The values observed for each of the pools were in good agreement among laboratories, and, as expected, according to currently accepted cutoff values, there would be no difficulty in separating smoker from nonsmoker samples in the observed results. All seven laboratories were capable of measuring the smoker-level samples in Pools C, D, and H, and five laboratories also reported results for the two pools with relatively high cotinine concentrations in the passive exposure range (pools B and G). Four of seven laboratories reported results for the moderate exposure Pool A, with a concentration of approximately 0.5 ng/ml, and three laboratories were able to measure all eight pools down to the lowest concentration level, which was approximately 0.05 ng/ml.

The mean, standard deviation, and range of values for each pool as measured by the participating laboratories are provided in Table 4. In general, there was excellent agreement among the laboratories for the pools that each laboratory could analyze. In

Table 5, the final 95% CIs for all pools with known spiked target values (A, B, C, and D) indicated that there was no significant bias at the 5% significance level among the laboratories reporting results for a pool. As recommended by ISO 5725-2, we calculated Mandel's h and k statistics by laboratory and pool and these results are given in Figure 1 with the 5% significance levels indicated on the charts with dotted lines. Only three pools exceeded the 5% limits in the h statistic plot, including one pool in Laboratory 1-B and two pools in Laboratory 3. Laboratories 3, 4, and 5 had a consistently negative bias relative to group means in Figure 1A, and Laboratory 7 was consistently positive. Such results might reflect minor calibration differences among the laboratories. Overall, however, the aggregate mean observed concentrations showed excellent agreement with the expected (target) concentrations in each case for the serum pools examined in this study.

In Figure 1B, the k statistic reflects the ratio of the within-laboratory standard deviation of each laboratory and pool to an averaged within-laboratory standard deviation of all seven laboratories for the given pool. All but two laboratories had results from one or two pools that showed significantly greater variability than the overall group variance for that pool, with the specific pools involved differing among the laboratories. Following recommendation 7.3.1.6 in ISO 5725-2:1994(E), we interchanged laboratory and pool in Figure 1B to examine if the pool outcomes were consistent with other laboratories (data not shown). This grouping indicated a mild potential concern for repeatability for Laboratory 1-A on fortified Pool A (the pool with the lowest target concentration) and Laboratory 7 on Pool D (the highest target pool). However, as indicated in Table 4, the relative standard deviations compiled by pool and laboratory remained reasonably low in all cases. Table 5 also provided assurance that the bias, when averaged over laboratories, was not significant for any of the fortified pools with defined target values.

Both the repeatability and the reproducibility deviations, as a function of mean values, showed a good fit as a line through the origin (data not shown). After the nonsignificant intercepts were removed, the models were $S_r = 0.034 \times m$ and $S_R = 0.079 \times m$, where m is the mean cotinine value ($R^2 > .98$ in both cases both before and after refitting). This indicates that the coefficient of variation was reasonably constant for all levels of m .

Discussion

The results from this study confirm an overall excellent level of performance for the seven laboratories participating in this study, which are all experienced laboratories currently analyzing serum cotinine. Among these laboratories, both GC/NPD with 5-methylcotinine as the internal standard and LC/MS/MS with isotopically labeled cotinine as the internal standard were shown to produce accurate and precise results within their calibration ranges. Not only all pools were ranked correctly by each laboratory but also the relative bias observed was minimal in each case.

All the participating laboratories were able to measure serum cotinine in samples with concentrations that might be associated with the active use of tobacco at about 3 ng/ml or higher (Benowitz, Bernert, Caraballo, Holiday, & Wang, 2009), and at least some of the laboratories were able to measure even

Interlaboratory comparability of serum cotinine measurements

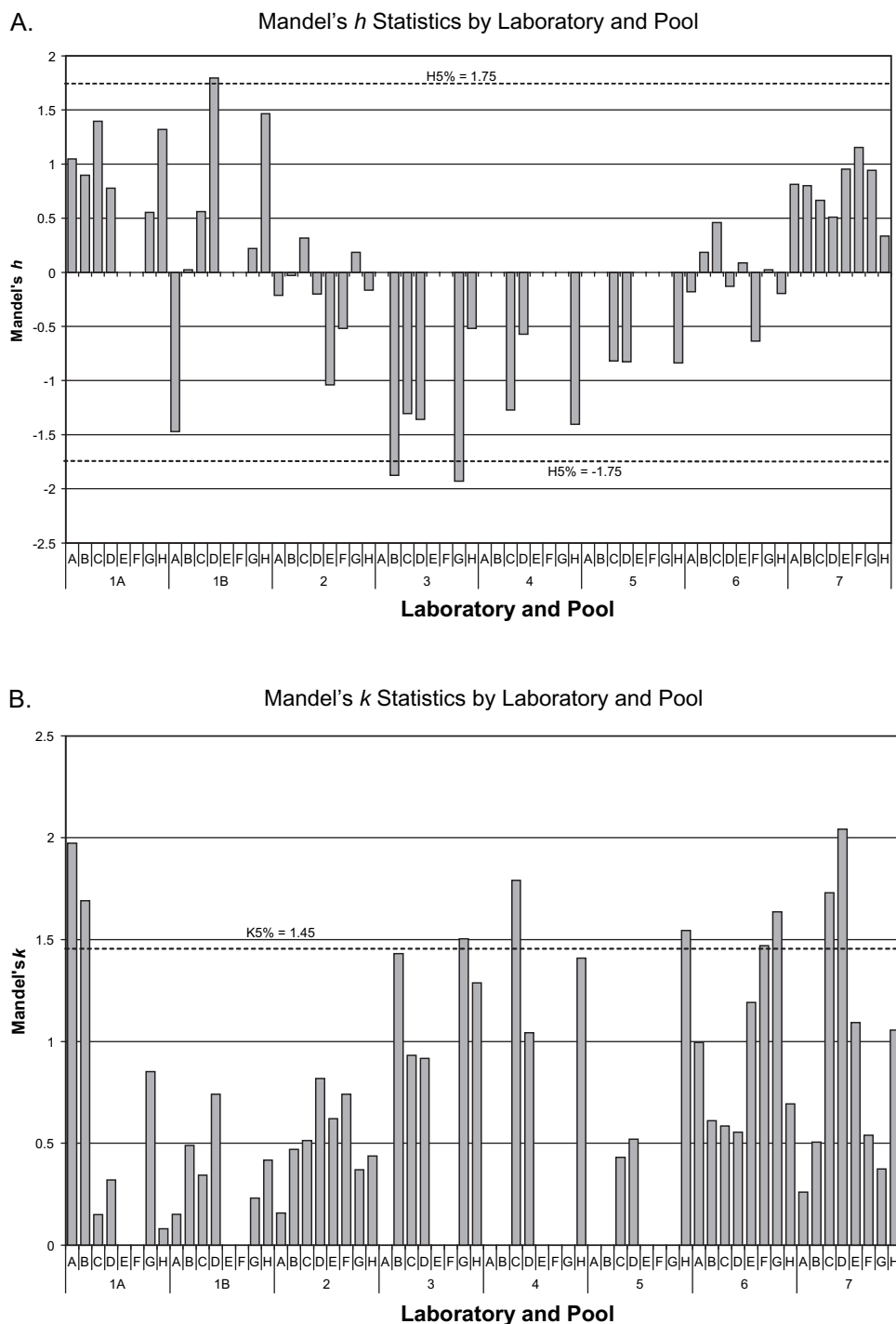


Figure 1. Mandel's h and k statistics.

the lowest concentration sample typically associated with minimal exposure to SHS with good precision and excellent agreement. The use of LC MS/MS may be of importance in this regard, since all three laboratories that were able to measure the lowest samples used some form of that technique for their assays.

Mandel's h and k statistic analyses were consistent with minimal bias for the fortified pools and with relatively minimum assay variance among the laboratories. Two laboratories

had very small but statistically significant biases for one or two pools in the study. Several laboratories had assay variances that were significantly greater than the pooled variance for one or occasionally two pools in the study; however, in only two cases were these differences of potential concern. In the case of Laboratory 1-A, this increased variance was associated with the lowest concentration fortified pool that was relatively close to the LOD for that laboratory's method; thus, greater variability might be expected. The reason for the relatively greater variance

in the other case (Pool D in Laboratory 7) was not apparent. However, as indicated in Table 4, even in that case, the relative standard deviation remained only 7.5%.

Assay sensitivity requirements may vary according to the intended applications. For studies involving the active use of tobacco, all the laboratories and methods used in this work were capable of excellent results. However, in the United States and other developed nations around the world, the capability of measuring even quite low concentrations of serum cotinine in nonsmokers is becoming more important as the incidence and intensity of exposure of nonsmokers to SHS decline. For example, in the third National Health and Nutrition Examination Survey, Phase 1, which covered the period 1988–1991, the median serum cotinine concentration for nonsmokers in the United States was 0.204 ng/ml among adults aged 20 years and older and 0.262 ng/ml for children aged 4–11 years. By 2001–2002, the median concentration among nonsmoker adults had declined to 0.034 ng/ml and it had declined to 0.067 ng/ml among children (Pirkle et al., 2006). When the incidence and intensity of exposure to SHS are high, a relatively high cutpoint is needed to exclude most nonsmokers. However, as the incidence and intensity of exposure to tobacco smoke continue to decline, the serum cotinine cutoff level between active smokers and nonsmokers exposed to SHS can also be expected to decline (Benowitz et al., 2009). Detecting and stratifying exposures at these declining concentration levels require sensitive and precise methodology.

The results of this investigation confirm that with the use of modern methodology, excellent agreement in measurements can be achieved among different laboratories. Thus, the greater variability in serum cotinine measurements by use of older technology that has been reported in the past (Biber et al., 1987) does not appear to be as much of a problem today; this decrease in analytical bias should facilitate the comparison of results from different studies and laboratories. The data from this investigation have also enabled us to assign consensus cotinine concentrations to the unfortified serum pools that we examined; these consensus concentrations will be of value in future method development and evaluations.

Funding

This work was supported by the Centers for Disease Control and Prevention (JTB, DBH, and CSS; 5U59EH223392-05 to KMA) and by the Flight Attendant Medical Research Institute and the National Institutes of Health (DA12393 to PJ and NLB).

Declaration of Interests

All authors state that they have no competing interests in this work.

Acknowledgments

We gratefully acknowledge the valuable assistance with statistical analyses provided by Jiantong Wang of RTI International. We also thank Tia Harmon, James R. Akins, LaQuasha Gaddis, and Ricky Alexander (Centers for Disease Control and Prevention); Lisa Yu and Minjiang Duan (University of California, San Francisco); James Daly (New York State Department of Health Wadsworth Center); Emma Tzioumis (Children's Hos-

pital Boston); and Amy Adler, Mary Dallman, and the Laboratory Staff, Toxicology and Drug Monitoring Laboratory, Mayo Clinic, for conducting measurements of serum cotinine as part of this study. The findings and conclusions in this report are those of the authors and do not necessarily represent the views of the Centers for Disease Control and Prevention. Use of trade names and commercial sources is for identification only and does not constitute endorsement by the U.S. Department of Health and Human Services or the Centers for Disease Control and Prevention.

References

- Benowitz, N. L. (1996). Cotinine as a biomarker of environmental tobacco smoke exposure. *Epidemiological Review*, 18, 188–204.
- Benowitz, N. L., Bernert, J. T., Caraballo, R. S., Holiday, D. B., & Wang, J. (2009). Optimal serum cotinine levels for distinguishing cigarette smokers and nonsmokers within different racial/ethnic groups in the United States between 1999 and 2004. *American Journal of Epidemiology*, 169, 236–248.
- Biber, A., Scherer, G., Hoepener, H., Adlkofer, F., Heller, W., Haddow, J., et al. (1987). Determination of nicotine and cotinine in human serum and urine: An interlaboratory study. *Toxicology Letters*, 35, 45–52.
- Borland, R., Mullins, R., Trotter, L., & White, V. (1999). Trends in environmental tobacco smoke restrictions in the home in Victoria, Australia. *Tobacco Control*, 8, 266–271.
- Centers for Disease Control and Prevention. (2007). Cigarette smoking among adults—United States, 2006. *Morbidity and Mortality Weekly Report* [serial online], 56, 1157–1161. Retrieved 8 November 2007, from <http://www.cdc.gov/mmwr/preview/mmwrhtml/mm5644a2.htm>
- Department of Health and Human Services. (2006). *The health consequences of involuntary exposure to tobacco smoke: A report of the Surgeon General*. Centers for Disease Control and Prevention, National Centers for Chronic Disease Prevention and Health Promotion, Office of Smoking and Health.
- Gilmore, J. (2002). *Report on smoking in Canada, 1985 to 2001* (Statistics Canada, Catalogue 82F0077XIE). Retrieved March 1, 2009, from <http://www.statcan.ca/english/research/82F0077XIE/82F0077XIE2001001.pdf>
- Hariharan, M., VanNoord, T., & Greden, J. F. (1988). A high-performance liquid chromatographic method for routine simultaneous determination of nicotine and cotinine in plasma. *Clinical Chemistry*, 34, 724–729.
- ISO 5725-1. (1994). *Accuracy (trueness and precision) of measurement methods and results: Part 1. General principles and definitions*. Geneva, Switzerland: International Organization for Standardization.
- ISO 5725-2. (1994). *Accuracy (trueness and precision) of measurement methods and results: Part 2. Basic method for the determination of repeatability and reproducibility of a standard measurement method*. Geneva, Switzerland: International Organization for Standardization.

Interlaboratory comparability of serum cotinine measurements

- ISO 5725-4. (1994). *Accuracy (trueness and precision) of measurement methods and results: Part 4. Basic methods for the determination of the trueness of a standard measurement method*. Geneva, Switzerland: International Organization for Standardization.
- Jacob, N., Berny, C., Boyer, J.-C., Capolaghi, B., de l'Homme, G., Desch, G., et al. (2005). Dosage urinaire de la cotinine et des metabolites de la nicotine. *Annales de Biologie Clinique*, 63, 397–409.
- Jarvis, M. J., Goddard, E., Higgins, V., Feyerabend, C., Bryant, A., & Cook, D. G. (2000). Childrens's exposure to passive smoking in England since the 1980s: Cotinine evidence from population surveys. *British Medical Journal*, 321, 343–345.
- Mackay, J., Eriksen, M., & Shafey, O. (2006). *The tobacco atlas* (2nd ed.). Atlanta, GA: American Cancer Society.
- Mandel, J. (1998). Interlaboratory study. In P. Armitage & T. Colton (Eds.), *Encyclopedia of biostatistics* (pp. 2073–2076). Hoboken, NJ: John Wiley & Sons.
- Pirkle, J. L., Bernert, J. T., Caudill, S. P., Sosnoff, C. S., & Pechacek, T. F. (2006). Trends in the exposure of nonsmokers in the U.S. population to secondhand smoke: 1988–2002. *Environmental Health Perspectives*, 114, 853–858.
- Strong, K., Guthold, R., Yang, J., Lee, D., Petit, P., & Fitzpatrick, C. (2008). Tobacco use in the European region. *European Journal of Cancer Prevention*, 17, 162–168.