Environment International 64 (2014) 129–133

Contents lists available at ScienceDirect

Environment International

journal homepage: www.elsevier.com/locate/envint

# Correction and comparability of phthalate metabolite measurements of Canadian biomonitoring studies (2007–2012)

Éric Langlois [a,*], Gurusankar Saravanabhavan [b], Tye E. Arbuckle [c], Suzelle Giroux [d]

[a] Centre de Toxicologie du Québec (CTQ), Institut National de Santé Publique du Québec (INSPQ), 945 Ave Wolfe, Québec, Québec G1V 5B3, Canada
[b] National Biomonitoring Section, Chemicals Surveillance Bureau, Environmental and Radiation Health Sciences Directorate, Healthy Environments and Consumer Safety Branch, Health Canada, 269 Laurier Ave. W., Ottawa, Ontario K1A 0K9, Canada
[c] Population Studies Division, Environmental Health Science and Research Bureau, Environmental and Radiation Health Sciences Directorate, Healthy Environments and Consumer Safety Branch, Health Canada, Environmental Health Center, 50 Colombine Dr., Tunney's Pasture, Ottawa, Ontario K1A 0K9, Canada
[d] Household Survey Methods Division, Statistics Canada, R.H. Coats Building, 100 Tunney's Pasture Driveway, Ottawa, Ontario K1A 0T6, Canada

ARTICLE INFO

*Article history:*
Received 6 August 2013
Accepted 4 December 2013
Available online 8 February 2014

*Keywords:*
Phthalate metabolites
Biomonitoring
Correction factor
Accuracy
Comparability
Standards

ABSTRACT

Phthalate metabolites are often measured in biomonitoring studies to evaluate a population's exposure to ubiquitous phthalates. During the course of national biomonitoring studies in Canada, we identified an issue with the accuracy of several commercial phthalate metabolite standards that are commonly used in such studies. The validity of the results from these studies was then questioned. Altogether, three (3) large studies were affected, involving a total of 9302 samples and 105 000 individual phthalate metabolite measurements. Data from our previous investigation suggested that the inaccuracies in the commercially-available phthalate metabolite standards were compound- and lot-specific. Therefore, an approach was developed to derive correction factors for each lot of phthalate metabolite standard and was applied to the previously-acquired measurements with the goal of obtaining accurate and comparable data. A statistical analysis was performed to support the approach. It is expected that the corrected phthalate metabolite data from all three Canadian biomonitoring studies are comparable to one another. However, caution is still advised when comparing data obtained from biomonitoring studies for which the calibration standards have not been investigated for their accuracy. Suggestions are made based on quality assurance aspects to improve the validity of phthalate metabolite measurements.



## 1. Introduction

Phthalates are a group of synthetic chemicals that have numerous industrial and commercial applications. The general public is exposed to these ubiquitous chemicals on a daily basis through the use of cosmetics, consumer products, drugs, building materials, PVC plastic materials, etc. Urinary phthalate metabolite measurements are frequently included in human biomonitoring (HBM) studies, such as NHANES (United States) (Centers for Disease Control and Prevention, 2012a; Silva et al., 2004), GerES (Germany) (Becker et al., 2009; Koch et al., 2007), CHMS (Canada) (Health Canada, 2013; Saravanabhavan et al., in press), MIREC (Canada) (Arbuckle et al., 2013), FNBI (Canada) (LaCorte and Wuttke, 2012) and ELFE (France) (Zeman et al., 2013). The data from such biomonitoring initiatives are used by scientists and policy makers to evaluate the magnitude of exposure and time trends, identify at-risk populations, establish reference ranges, identify knowledge gaps and develop relevant scientific policies. To properly interpret biomonitoring data, they must be accurate, comparable and acquired using a fully validated analytical method.

There are only a few published reports on comparability issues of HBM data related to the accuracy of laboratory measurements in the literature. For example, the CDC discovered that a drift in the vitamin D status between two consecutive NHANES surveys was not related to a real change in the population but rather to a reformulation of the manufacturer calibrators (Centers for Disease Control and Prevention, 2010; Yetley et al., 2010). This change affected the accuracy of the vitamin D data from 1988 to 1994 and therefore the comparability and interpretability of this NHANES dataset. Other drifts in the vitamin D assay performance were also observed in NHANES 2003–2006 and related to a calibration lot change. The CDC adjusted the data to account for those

*Abbreviations:* PVC, polyvinyl chloride; NHANES, National Health and Nutrition Examination Survey; GerES, German Environmental Survey; FNBI, First Nations Biomonitoring Initiative; ELFE, Etude Longitudinale Française depuis l'Enfance; CDC, Centers for Disease Control and Prevention; Me-PFOSA-AcOH, 2-(N-methyl-perfluorooctane sulfonamido)acetic acid; CHMS, Canadian Health Measures Survey; MIREC, Maternal–Infant Research on Environmental Chemicals; P4, Plastics and Personal care Products use in Pregnancy study; UPLC–MS-MS, ultrahigh pressure liquid chromatography–tandem mass spectrometry; CIL, Cambridge isotope laboratories; MMP, monomethyl phthalate; MEP, monoethyl phthalate; MnBP, mono-n-butyl phthalate; MiBP, mono-isobutyl phthalate; MBzP, monobenzyl phthalate; MCHP, monocyclohexyl phthalate; MEHP, mono-2-ethylhexyl phthalate; MEHHP, mono-(2-ethyl-5-hydroxyhexyl) phthalate; MEOHP, mono-(2-ethyl-5-oxohexyl) phthalate; MECPP, mono-(5-carboxy-2-ethylpentyl) phthalate; MOP, mono-n-octyl phthalate; MCPP, mono-3-carboxylpropyl phthalate; MNP, mono-isononyl phthalate; GLM, generalized linear model; NIST, National Institute of Standards and Technology; PT, proficiency testing; SRM, standard reference material; CRM, certified reference material.

* Corresponding author. Tel.: +1 418 650 5115x4644.
*E-mail address:* eric.langlois@inspq.qc.ca (É. Langlois).

0160-4120/$ – see front matter © 2013 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.envint.2013.12.002

two issues as described in an analytical note released in 2010 (Centers for Disease Control and Prevention, 2010). Several papers also described comparability issues with folate and vitamin B-12 datasets in 1991–1994 and 1999–2006 NHANES surveys (Bock and Eckfeldt, 2011; Yetley and Johnson, 2011; Yetley et al., 2011a, 2011b, 2011c). The CDC also reported a purity issue with a perfluorinated compound calibration standard used in their projects since the early 2000s. As a result, the CDC applied a correction factor to a decade of Me-PFOSA-AcOH measurements from the NHANES survey (Centers for Disease Control and Prevention, 2012b). While the CDC was responsive and thorough in reporting these issues, the inaccuracies with commercial calibration standards are likely not limited to the CDC studies. If not well described and disseminated, such issues may remain undisclosed, and the management and interpretation of the biomonitoring data may not be performed appropriately.

The Centre de Toxicologie du Québec (CTQ), at the Institut National de Santé Publique du Québec (INSPQ), was responsible for the analysis of urinary phthalate metabolites in three Canadian HBM studies described in Section 2.1. Overall, CTQ analyzed 9302 samples for a total of 105 000 individual phthalate metabolite measurements in these studies. During the course of the work, CTQ identified an issue with the accuracy of the commercial phthalate metabolite standards that were used as calibrators. Because any inaccuracy in the calibration standards would adversely affect the quality of the data and limit the comparability of the results between laboratories, we carried out a comprehensive assessment of the accuracy of these standards (Langlois et al., 2012). The inaccuracies in the "certified" concentrations of the standards, up to 76% (Table 1), were observed to be compound-dependent, form-dependent, supplier-dependent and lot-dependent. Moreover, CTQ used calibration standards from different lots during the course of the three biomonitoring studies. Therefore, the accuracy issue also limited the comparability of the phthalate metabolite data from within these studies, in spite of being from the same laboratory using the same analytical method. Therefore, there was a need to correct the data using lot-specific correction factors to improve the accuracy and comparability of the measurements both within and between laboratories. Even though correction factors are a last resort, they remain particularly useful in situations where re-analysis is impractical and/or too expensive.

When informed of this inaccuracy issue, the CDC carried out their own investigation and reported inaccuracies with 5 of their phthalate metabolite standards (Centers for Disease Control and Prevention, 2012a). The CDC also developed compound-specific correction factors and used them to adjust the data collected since the 1999–2000 NHANES survey period. As previously expected, the inaccuracy issue with several phthalate metabolite standards appears to be widespread (Langlois et al., 2012).

Hence, in addition to bringing such an issue to light, it appears to be particularly important to present a solution for laboratories that analyze phthalate metabolites and for HBM managers and data users who may be faced with this issue.

This paper describes an approach for the development of compound/lot-specific correction factors to adjust phthalate metabolite data from the aforementioned Canadian biomonitoring studies. To support the approach, a statistical analysis of the calibration data is also presented. Finally, we discuss some aspects of comparability and present suggestions of improvements related to quality assurance.

## 2. Material and methods

### 2.1. Description of the Canadian human biomonitoring studies

#### 2.1.1. Canadian Health Measures Survey (CHMS)

The CHMS, first launched in 2007, is an on-going comprehensive direct health measures survey in Canada with a goal to collect important data on the health of Canadians to fill the existing knowledge gaps. This biennial survey collects blood and urine samples to test for chronic and infectious diseases, nutrition and environment markers among other things (Statistics Canada, 2007). The CHMS (2007–2009) covers the population 6 to 79 years of age whereas in CHMS (2009–2011) 3 to 5 years were also included. Urinary phthalate metabolites were measured in a subsample of 3237 people 6 to 49 years of age in CHMS 2007–2009 and on 2563 people 3 to 79 years of age in CHMS 2009–2011.

#### 2.1.2. Maternal–Infant Research on Environmental Chemicals (MIREC)

MIREC is a national-level research study that recruited 2000 pregnant women from 10 major Canadian cities (Arbuckle et al., 2013). Women were followed through pregnancy and for up to ten weeks after birth. One of the goals of the study was to obtain a national profile of in utero and lactational exposure to environmental contaminants. Biological markers of environmental chemical exposure were measured in the mothers' blood, urine, hair, and breast milk and in their babies' umbilical cord blood and meconium.

#### 2.1.3. Plastics and personal-care products use in pregnancy (P4)

The P4 study recruited 80 pregnant women, followed them through pregnancy and post-delivery and collected multiple maternal urine samples, detailed consumer product/food packaging diaries, infant urine and meconium, and breast milk. The objectives of this study were to examine temporal trends in exposure over a 24-hour period and during pregnancy and to evaluate the validity of a spot urine sample to estimate an individual's exposure. Biospecimens were analyzed for phthalate metabolites, bisphenol A, triclosan and triclocarban.

### 2.2. Analysis of urine samples

All urine sample analyses for the three HBM studies were carried out by UPLC–MS-MS at the Centre de Toxicologie du Québec, INSPQ using a

**Table 1**
Relative inaccuracies (%) of phthalate metabolite standards against CIL-neat standards.

| Form | Set of standards (supplier) | MMP | MEP | MiBP | MnBP | MCHP | MBzP | MEHP | MEHHP | MEOHP | MOP | MCPP | MNP |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Certified solutions | CIL-2006 | [-25%] | -2% | - | [-47%] | -1% | [-63%] | [-29%] | -11% | -7% | 12% | [-22%] | [-39%] |
| | CIL-2009 | -9% | -11% | -5% | -13% | -12% | [-76%] | -14% | [-21%] | 4% | 2% | [-70%] | [-16%] |
| Neat | CanSyn | 9% | -5% | 4% | -1% | -6% | -1% | -8% | -6% | -4% | 10% | -7% | -3% |
| | CIL | 0% | 0% | 0% | 0% | 0% | 0% | 0% | 0% | 0% | 0% | 0% | 0% |
| | TRC | 7% | -3% | 4% | -6% | -10% | -2% | -11% | 9% | 9% | 12% | -2% | -6% |

–: not applicable.
Values greater than 15% are in boxes.
Partial reproduction from (Langlois et al., 2012).

method described previously (Langlois et al., 2012) under an ISO-17025 accredited quality system.

Overall, the following 12 metabolites were analyzed: MMP, MEP, MiBP, MnBP, MBzP, MCHP, MEHP, MEHHP, MEOHP, MOP, MCPP, and MNP. All standards were obtained from Cambridge Isotope Laboratories (CIL). The corresponding $^{13}C_4$-isotope labeled analogs of the phthalate metabolites were used as internal standards.

Two sets of standards from different lots (CIL-2006 and CIL-2009) were used for the analysis of the urine specimens. The specific lot numbers associated with each standard of these two sets have been previously provided (Langlois et al., 2012). The standards of the CIL-2006 set were used in the analysis of the urine samples from the CHMS 2007–2009 and MIREC studies. Those of the CIL-2009 set were used in the analysis of the urine samples from the CHMS 2009–2011 and P4 studies. MiBP was included in the CHMS 2009–2011 study only.

### 2.3. Calibration standards analysis

In addition to the two sets of certified solution standards mentioned above, a set of neat standards of the same analytes was also used (CIL-neat) in the calibration standards analysis. These standards were all from CIL. Individual stock solutions of each standard were prepared in acetonitrile at 100 μg/mL.

Working solution mixes of 2000, 200 and 20 ng/mL were prepared by diluting appropriate volumes of individual stock solutions with demineralized water. These working solution mixes were used to prepare the calibration points along with a labeled isotope standard solution mix used as internal standard.

A calibration curve was prepared for each single lot of standards from the CIL-2006, CIL-2009 and CIL-neat sets. Each calibration curve comprised eight calibration concentrations in triplicate in the range of 0.5 ng/mL to 200 ng/mL for a total of 24 calibration points. The MEP calibration curves comprised three additional calibration concentrations in triplicate from 200 ng/mL to 2000 ng/mL for a total of 33 calibration points.

Instrumental analysis was carried out for each lot of standards from the CIL-2006, CIL-2009 and CIL-neat sets using the method mentioned above. The calibration curve data were used for the statistical analysis described below.

### 2.4. Statistical method

To assess the need to derive correction factors for a given metabolite, the slopes were tested for significant differences. A linear regression was done to derive the slopes of the calibration curves for each of the phthalate metabolites. The dependent variable was the analytical calibration response obtained from the standards analysis described in Section 2.3 and the independent variables were the purported concentration, the lot of standard used for the measure (CIL-2006, CIL-2009 and CIL-neat) and the interaction term between the two. Contrasts were used to statistically compare the slopes of the three lots of standards of each phthalate metabolite (CIL-2006, CIL-2009 and CIL-neat). Statistical differences were evaluated at the 95% and 99% levels. The statistical analyses were carried out using the GLM procedure of the SAS 9.2v software.

### 2.5. Development of correction factors

For each single lot of standards from CIL-2006 and CIL-2009, the slopes of the corresponding calibration curves described above were used to derive the correction factors using the following equations:

$$F_{CIL-2006} = \frac{m_{CIL-2006}}{m_{CIL\ \ neat}} \quad (1)$$

$$F_{CIL-2009} = \frac{m_{CIL-2009}}{m_{CIL\ \ neat}} \quad (2)$$

where $F$ is the correction factor, and $m$ is the slope.

### 2.6. Revised concentrations

The revised data of the original measurements were obtained by multiplying the original measured concentration of each phthalate metabolite by the correction factor of the corresponding standard, as presented by the following equations:

$$C_{revised,CHMS\,2007-2009} = F_{CIL-2006} C_{original,CHMS-1} \quad (3)$$

$$C_{revised,MIREC} = F_{CIL-2006} C_{original,MIREC} \quad (4)$$

$$C_{revised,CHMS\,2009-2011} = F_{CIL-2009} C_{original,CHMS-2} \quad (5)$$

$$C_{revised,P4} = F_{CIL-2009} C_{original,P4} \quad (6)$$

where $C_{revised}$ is the revised concentration, and $C_{original}$ is the original concentration.

## 3. Results and discussion

The "certified" concentrations of several commercial standard solutions used in the urinary phthalate metabolite measurements in three large-scale HBM studies in Canada were found to be inaccurate by up to 76% (Table 1). The time and cost involved in re-running the biological samples from these HBM studies would have been prohibitive. Therefore, we developed an approach to derive compound-specific and lot-specific correction factors to adjust the phthalate metabolite data previously obtained from these HBM studies.

To obtain scientifically defensible correction factors, the use of a "zero-bias" reference is necessary. A standard reference material (SRM, from NIST) or a certified reference material (CRM, from any reliable source) would have been an appropriate basis from which to derive the correction factors. However, to the best of our knowledge, such reference materials for the analysis of phthalate metabolites in human biological matrices were unavailable at the time of this work.

Therefore, our approach was based on the results from our previously published accuracy investigation study in which we found that the regression parameters of the calibration curves developed with neat standards from three different suppliers were comparable to one another, suggesting that these sets of standards are accurate (we refer the reader to our previous article for more details (Langlois et al., 2012)). These three sets of neat standards may therefore be considered appropriate for the derivation of the correction factors. We chose the standard set from CIL (CIL-neat) to derive appropriate and compound-specific correction factors. To allow comparability with future measurements, the standards from this set were also chosen as the usual calibration standards at the CTQ for future phthalate metabolite analyses.

### 3.1. Statistical analysis

In urine specimens, the regression equations of the calibration curves were used to quantify phthalate metabolite concentrations (Section 2.2). Therefore, to assess the need to derive correction factors for a given metabolite, we examined the significance of the deviations between the regression slopes of the calibration curves obtained using the standards from the CIL-2006 and CIL-2009 sets and the corresponding neat reference standard of the CIL-neat set, as described in Section 2.3. In Table 2, the statistical tests performed on each lot of phthalate metabolite standards indicated that the slopes were significantly different ($p < 0.05$) for almost all phthalate metabolites produced using CIL-2006 and CIL-2009 compared to those obtained using

**Table 2**
Calibration curve slopes of each standard of the CIL-2006 and CIL-2009 sets of certified solution standards and their respective p-values and correction factors, along with the corresponding calibration curve slopes of the CIL-neat set of standards.

| Analytes | CIL-neat | CIL-2006 | | | CIL-2009 | | |
|---|---|---|---|---|---|---|---|
| | Calibration curve slope | Calibration curve slope | p-Value for comparison to CIL-neat | Correction factor | Calibration curve slope | p-Value for comparison to CIL-neat | Correction factor |
| MMP | 0.07526 | 0.05611 | <0.0001 | 0.75 | 0.06863 | <0.0001 | 0.91 |
| MEP | 0.02449 | 0.02408 | 0.0195 | 0.98 | 0.02172 | <0.0001 | 0.89 |
| MiBP | 0.05789 | – | – | – | * | – | * |
| MnBP | 0.05694 | 0.03039 | <0.0001 | 0.53 | 0.04949 | <0.0001 | 0.87 |
| MCHP | 0.05111 | 0.05042 | 0.1001 | 0.99 | 0.04513 | <0.0001 | 0.88 |
| MBzP | 0.06536 | 0.02387 | <0.0001 | 0.37 | 0.01588 | <0.0001 | 0.24 |
| MEHP | 0.10240 | 0.07267 | <0.0001 | 0.71 | 0.08826 | <0.0001 | 0.86 |
| MEHHP | 0.07646 | 0.06841 | <0.0001 | 0.89 | 0.06074 | <0.0001 | 0.79 |
| MEOHP | 0.06442 | 0.06004 | <0.0001 | 0.93 | 0.06720 | <0.0001 | 1.04 |
| MOP | 0.06884 | 0.07687 | <0.0001 | 1.12 | 0.07013 | 0.0635 | 1.02 |
| MCPP | 0.05899 | 0.04630 | <0.0001 | 0.78 | 0.01783 | <0.0001 | 0.30 |
| MNP | 0.07100 | 0.04362 | <0.0001 | 0.61 | 0.05981 | <0.0001 | 0.84 |

–: Not applicable.
*: For MiBP, the calibration curve was processed using a quadratic equation. Therefore, the corrected concentrations were derived using the following equation:

$$C_{revised} = 1.36 \times 10^{-3}(C_{CIL-2009})^2 + 0.951(C_{CIL-2009}) - 0.208$$

CIL-neat. This suggests that the measurements obtained using the CIL-2006 and the CIL-2009 sets as calibration standards would produce erroneous results, necessitating the development and application of correction factors in order to obtain accurate data. The two exceptions were the MCHP standard from CIL-2006 and the MOP standard from CIL-2009 whose slopes were not statistically different from the slope of their corresponding CIL-neat standard counterpart.

### 3.2. Correction factors and correction of the measurements

Correction factors for each phthalate metabolite from the CIL-2006 and CIL-2009 sets were calculated using Eqs. (1) and (2) and are presented in Table 2. In total, 24 compound- and lot-specific correction factors were obtained. For CIL-2006, the correction factors ranged from 0.37 (MBzP) to 1.12 (MOP). In the case of CIL-2009, the correction factors ranged from 0.24 (MBzP) to 1.04 (MEOHP). In the case of the MiBP standard from CIL-2009, the calibration curve showed a better fit with a quadratic equation rather than a linear equation. Hence, the correction factor associated with this standard was a quadratic equation instead of a simple multiplying factor.

It should be noted that, for standards whose correction factors are near 1, the standard possessed comparable accuracy to that of the neat reference standard, and correcting the results may not be necessary. However, to maintain consistency, all 105 000 individual phthalate metabolite measurements were corrected using the appropriate correction factors shown in Table 2.

The correction factors obtained for the CIL-2006 standards were used along with Eqs. (3) and (4) to correct the urinary phthalate metabolite measurement data from the CHMS (2007–2009) (Health Canada, 2013; Saravanabhavan et al., in press) and MIREC studies. The correction factors obtained for the CIL-2009 standards were applied to the urinary phthalate metabolite measurements of the CHMS (2009–2011) (Health Canada, 2013) and P4 studies, along with Eqs. (5) and (6). The limits of detection and the limits of quantification were also adjusted using the appropriate correction factors. The data from MIREC and P4 will be released in the near future.

It is worth reminding here that the different lots of certified solution standards from any supplier are likely to possess different inaccuracies, so different correction factors will need to be developed and applied. It should then be emphasized that each laboratory facing this accuracy issue should derive its own correction factors if its lots of certified solution standards are different than those listed in Table 2 of our previous article (Langlois et al., 2012). It is also advisable that the laboratories do not limit their investigation to the analytes presented here but to their own full line of phthalate metabolite standards, as it is likely that this issue may extend to other phthalate metabolites. The several inconveniences related to the correction of a large number of previous measurements underscore the importance, for any laboratory, of performing such verification before the beginning of a study.

### 3.3. Accuracy and comparability

The use of accurate standards or well-defined correction factors is essential to the release of accurate and comparable data. We believe that after adjustment with the correction factors developed in this work, the phthalate metabolite data from the aforementioned Canadian HBM studies are comparable to one another. It is also expected that the corrected data and any future phthalate metabolite measurements from CTQ/INSPQ will not suffer from within-laboratory comparability issues related to the standard used, considering the use of accurate neat standards as regular calibration standards from now on. The release of accurate data will favor the proper use and better interpretation of these data.

It is also expected that the revision of the phthalate metabolite data from the CHMS, MIREC, and P4 studies will favor comparability to those of the NHANES because these data were all corrected by the CTQ/INSPQ

and the CDC following the respective accuracy investigations they performed. However, usual cautions are still applicable because other variables, such as the study designs, may affect the comparability.

Other than the CDC, we are not aware of any other laboratory or biomonitoring studies that acknowledged inaccuracies in their phthalate metabolite data. However, it is very likely that urinary phthalate metabolite measurements from other laboratories worldwide are affected by similar inaccuracy issues. Obviously, this may cause comparability issues. Therefore, any comparisons with and within other studies should be exercised with caution and should be restricted to data acquired with standards of equal accuracy. In the objective of improving the accuracy and the comparability of upcoming phthalate metabolite measurements from any studies worldwide, Langlois et al. proposed several solutions that may be considered by laboratories, study managers and suppliers.

### 3.4. Quality assurance

It is commonly assumed that commercial calibration standards are accurate, particularly the certified standards. However, as our results illustrate, such assumptions must be thoroughly checked to avoid producing biased measurements. From the perspective of improvement and to favor the identification of a such issue, we present the following recommendations focusing on quality assurance:

Firstly, to make available of a proficiency testing (PT) scheme offering a complete line of phthalate metabolites with accurate reference values. Such scheme would provide laboratories with a complete and unbiased evaluation of the accuracy and the comparability of their data. In particular, the accuracy of the reference values should be demonstrated in regard to the actual accuracy issue.

The availability of a CRM/SRM, including a complete line of phthalate metabolites in urine, is also highly desirable. Such reference material may be used to monitor the accuracy of analytical results. The development of a reference material has been previously recognized and implemented to resolve an accuracy issue related to vitamin D analysis (de la Hunty et al., 2010; Phinney et al., 2012; Yetley et al., 2010). The availability of a CRM/SRM would also be beneficial in estimating correction factors in the case of previous measurements for which the calibration standard lots used are no longer available.

## 4. Conclusions

Following our previous findings on the inaccuracy of phthalate metabolite standards, we reported a method to derive compound- and lot-specific correction factors from the calibration data. The statistical analysis performed on the calibration data suggested that correction of the analytical results was necessary to obtain accurate and comparable data. The correction factors were applied to 105 000 previously acquired urinary measurements from three Canadian HBM studies. Even though the corrected data of those studies are now considered accurate, any comparisons with and within other worldwide studies should be exercised with caution and should be restricted to data acquired with standards of equal accuracy. We still recommend caution in the comparison of any data produced from standards that have not been investigated for their accuracy. We believe that the approach presented here will be beneficial for laboratories involved in phthalate metabolite analysis to test, identify and adjust for inaccuracies in the commercial solution standards of phthalate metabolites. This will increase the accuracy, comparability and interpretation of the available data and favor a better use of these data. From a perspective of prevention and improvement, we also suggested the development of a CRM/SRM for urinary phthalate metabolite measurements.

## Acknowledgments

We would like to express our kind thanks to Pierre Ayotte from the Institut National de Santé Publique du Québec for reviewing the manuscript of this article. This work was carried out with financial support from CTQ/INSPQ, Health Canada and Statistics Canada.

## References

Arbuckle TE, Fraser WD, Fisher M, Davis K, Liang CL, Lupien N, et al. Cohort profile: the maternal–infant research on environmental chemicals research platform. Paediatr Perinat Epidemiol 2013;27:415–25.
Becker K, Goen T, Seiwert M, Conrad A, Pick-Fuss H, Muller J, et al. GerES IV: phthalate metabolites and bisphenol A in urine of German children. Int J Hyg Environ Health 2009;212:685–92.
Bock JL, Eckfeldt JH. Advances in standardization of laboratory measurement procedures: implications for measuring biomarkers of folate and vitamin B-12 status in NHANES. Am J Clin Nutr 2011;94:332S–6S.
Centers for Disease Control and Prevention. Revised analytical note for NHANES 2000–2006 and NHANES III (1988–1994) 25-hydroxyvitamin D analysis (Revised November 2010); 2010.
Centers for Disease Control and Prevention. Fourth national report on human exposure to environmental chemicals, updated tables, February 2012. CDC; 2012a.
Centers for Disease Control and Prevention. Fourth national report on human exposure to environmental chemicals, updated tables, September 2012. CDC; 2012b.
de la Hunty A, Wallace AM, Gibson S, Viljakainen H, Lamberg-Allardt C, Ashwell M. UK Food Standards Agency Workshop Consensus Report: the choice of method for measuring 25-hydroxyvitamin D to estimate vitamin D status for the UK National Diet and Nutrition Survey. Br J Nutr 2010;104:612–9.
Health Canada. Second report on human biomonitoring of environmental chemicals in Canada; 2013.
Koch HM, Becker K, Wittassek M, Seiwert M, Angerer J, Kolossa-Gehring M. Di-n-butylphthalate and butylbenzylphthalate — urinary metabolite levels and estimated daily intakes: pilot study for the German Environmental Survey on children. J Expo Sci Environ Epidemiol 2007;17:378–87.
LaCorte E, Wuttke S. The First Nations Biomonitoring Initiative—FNBI. Int J Hyg Environ Health 2012;215:168–71.
Langlois E, Leblanc A, Simard Y, Thellen C. Accuracy investigation of phthalate metabolite standards. J Anal Toxicol 2012;36:270–9.
Phinney KW, Bedner M, Tai SS, Vamathevan VV, Sander LC, Sharpless KE, et al. Development and certification of a standard reference material for vitamin D metabolites in human serum. Anal Chem 2012;84:956–62.
Saravanabhavan G, Guay M, Langlois E, Giroux S, Murray J, Haines D. Biomonitoring of phthalate metabolites in the Canadian population through the Canadian Health Measures Survey (2007–2009). Int J Hyg Environ Health 2013;216:652–61.
Silva MJ, Barr DB, Reidy JA, Malek NA, Hodge CC, Caudill SP, et al. Urinary levels of seven phthalate metabolites in the U.S. population from the National Health and Nutrition Examination Survey (NHANES) 1999–2000. Environ Health Perspect 2004;112:331–8.
Statistics Canada. Background papers on the Canadian Health Measures Survey. Health Reports Catalogue 82-003; 2007 (supplement).
Yetley EA, Johnson CL. Folate and vitamin B-12 biomarkers in NHANES: history of their measurement and use. Am J Clin Nutr 2011;94:322S–31S.
Yetley EA, Pfeiffer CM, Schleicher RL, Phinney KW, Lacher DA, Christakos S, et al. NHANES monitoring of serum 25-hydroxyvitamin D: a roundtable summary. J Nutr 2010;140:2030S–45S.
Yetley EA, Coates PM, Johnson CL. Overview of a roundtable on NHANES monitoring of biomarkers of folate and vitamin B-12 status: measurement procedure issues. Am J Clin Nutr 2011a;94:297S–302S.
Yetley EA, Pfeiffer CM, Phinney KW, Bailey RL, Blackmore S, Bock JL, et al. Biomarkers of vitamin B-12 status in NHANES: a roundtable summary. Am J Clin Nutr 2011b;94:313S–21S.
Yetley EA, Pfeiffer CM, Phinney KW, Fazili Z, Lacher DA, Bailey RL, et al. Biomarkers of folate status in NHANES: a roundtable summary. Am J Clin Nutr 2011c;94:303S–12S.
Zeman FA, Boudet C, Tack K, Floch BA, Brochot C, Pery AR, et al. Exposure assessment of phthalates in French pregnant women: results of the ELFE pilot study. Int J Hyg Environ Health 2013;216:271–9.