HAL
open science

# Within-Day, Between-Day, and Between-Week Variability of Urinary Concentrations of Phenol Biomarkers in Pregnant Women

Céline Vernet, Claire Philippat, Antonia Calafat, Xiaoyun Ye, Sarah Lyon-Caen, Valérie Siroux, Enrique Schisterman, Rémy Slama

## To cite this version:

Céline Vernet, Claire Philippat, Antonia Calafat, Xiaoyun Ye, Sarah Lyon-Caen, et al.. Within-Day, Between-Day, and Between-Week Variability of Urinary Concentrations of Phenol Biomarkers in Pregnant Women. Environmental Health Perspectives, 2018, 126 (3), pp.037005. 10.1289/EHP1994 . inserm-03156806

**HAL Id: inserm-03156806**

**https://inserm.hal.science/inserm-03156806v1**

Submitted on 2 Mar 2021

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

A Section 508–conformant HTML version of this article is available at https://doi.org/10.1289/EHP1994.

# Within-Day, Between-Day, and Between-Week Variability of Urinary Concentrations of Phenol Biomarkers in Pregnant Women

*Céline Vernet,[1] Claire Philippat,[1] Antonia M. Calafat,[2] Xiaoyun Ye,[2] Sarah Lyon-Caen,[1] Valérie Siroux,[1] Enrique F. Schisterman,[3] and Rémy Slama[1]*

[1]Team of Environmental Epidemiology applied to Reproduction and Respiratory Health, U1209, Inserm (Institut national de la santé et de la recherché médicale), IAB (Institute for Advanced Biosciences), CNRS (Centre national de la recherche scientifique), University Grenoble Alpes, Grenoble, France
[2]Organic Analytical Toxicology Branch, Division of Laboratory Sciences, Centers for Disease Control and Prevention, Atlanta, Georgia, USA
[3]Epidemiology Branch, Division of Intramural Population Health Research, Eunice Kennedy Shriver National Institute of Child Health and Human Development, Rockville, Maryland, USA

**Background:** Toxicology studies have shown adverse effects of developmental exposure to industrial phenols. Evaluation in humans is challenged by potentially marked within-subject variability of phenol biomarkers in pregnant women, which is poorly characterized.

**Objectives:** We aimed to characterize within-day, between-day, and between-week variability of phenol urinary biomarker concentrations during pregnancy.

**Methods:** In eight French pregnant women, we collected all urine voids over a 1-wk period (average, 60 samples per week per woman) at three occasions ($15 \pm 2$, $24 \pm 2$, and $32 \pm 1$ gestational weeks) in 2012–2013. Aliquots of each day and of the whole week were pooled within-subject. We assayed concentrations of 10 phenols in these pools, and, for two women, in all spot (unpooled) samples collected during a 1-wk period. We characterized variability using intraclass correlation coefficients (ICCs) with spot samples (within-day variability), daily pools (between-day variability), and weekly pools (between-week variability).

**Results:** For most biomarkers, the within-day variability was high (ICCs between 0.03 and 0.50). The between-day variability, based on samples pooled within each day, was much lower, with ICCs >0.60 except for bisphenol S (0.14, 95% confidence interval [CI]: 0.00, 0.39). The between-week variability differed between compounds, with triclosan and bisphenol S having the lowest ICCs (<0.3) and 2,5-dichlorophenol the highest (ICC >0.9).

**Conclusion:** During pregnancy, phenol biomarkers showed a strong within-day variability, while the variability between days of a given week was more limited. One biospecimen is not enough to efficiently characterize exposure; collecting biospecimens during a single week may be enough to represent well the whole pregnancy exposure for some but not all phenols. https://doi.org/10.1289/EHP1994

## Introduction

Phenols include high-production-volume chemicals with widespread uses in daily life products. For example, bisphenols are employed in the manufacture of epoxy resins and certain polymer plastics used in food and beverage containers and in other consumer products (Chen et al. 2016; INERIS 2014; NTP 2008). Parabens are used as preservatives in cosmetics, food, beverages, and pharmaceuticals; benzophenone-3, an ultraviolet-filter, is used in plastics and cosmetics; triclosan is used for its antibacterial properties in personal care products, clothing, or kitchenware (Cosmetic Ingredient Review Expert Panel 2008; Krause et al. 2012; NLM 2016); 2,4-dichlorophenol is used in the production of certain pesticides; and 2,5-dichlorophenol is a major metabolite of 1,4-dichlorobenzene, which is used in moth balls and room deodorizers (Crinnion 2010; HSDB 2016).

Concern exists regarding the health effects of phenols, which are potential endocrine disruptors, particularly following exposure during fetal life (Braun 2017). In terms of study design, most biomarker-based studies in humans relied on biomarker concentrations assessed in very few (one to three) spot biospecimens per pregnant woman. For chemicals with strong within-subject temporal variations, relying on a small number of biospecimens is expected to imperfectly characterize the average exposure (e.g., over a day, a week or longer), to lead to exposure misclassification, and consequently bias dose–response functions (Carroll et al. 2006; Perrier et al. 2016). The biological half-life of phenols in pregnant women is not known and could strongly differ from that of nonpregnant women, for example, as is the case for urinary biomarkers of tobacco smoke exposure such as cotinine, which has been found to have about twice as fast elimination half-life during pregnancy compared with postpartum (Dempsey et al. 2002). Studies based on nonpregnant adults reported a short (<12-h) half-life for some phenols (Janjua et al. 2007; Sandborgh-Englund et al. 2006; Völkel et al. 2002). Consequently, the relevance of relying on one spot biospecimen to provide a proxy of exposure for time windows of one day or longer is probably limited. This issue is of importance given the expected impact of exposure misclassification on bias in dose–response functions relating biomarker levels to health parameters (Perrier et al. 2016). Several studies evaluated the reproducibility of urinary phenol concentrations during pregnancy (Bertelsen et al. 2014; Braun et al. 2011, 2012; Guidry et al. 2015; Jusko et al. 2014; Meeker et al. 2013; Philippat et al. 2013; Smith et al. 2012; Stacy et al. 2017). These studies relied on generally two or three spot biospecimens collected from each pregnant woman several weeks or months apart. Such a design did not allow characterizing the within-day or the within-week variability in biomarkers concentrations. Based on complete urine collections throughout several days in eight nonpregnant participants, two studies reported high within-subject and between-day variability for bisphenol A (Koch et al. 2014; Ye et al. 2011), whereas this variability was relatively small for some parabens, triclosan, and benzophenone-3 (Koch et al. 2014). High within-day variability of bisphenol A concentrations was also reported in 66 pregnant women with complete urine

Address correspondence to C. Vernet, Institute for Advanced Biosciences, Centre de Recherche INSERM-CNRS-Université Grenoble-Alpes U1209, UJF Site Santé, BP 170, La Tronche, Grenoble, 38042 Cedex 9, France. Telephone: 33 4 76 54 94 21. Email: celine.vernet@univ-grenoble-alpes.fr

Supplemental Material is available online (https://doi.org/10.1289/EHP1994).

The authors declare they have no actual or potential competing financial interests.

Received 5 April 2017; Revised 12 December 2017; Accepted 12 December 2017; Published 16 March 2018.

**Note to readers with disabilities:** *EHP* strives to ensure that all journal content is accessible to all readers. However, some figures and Supplemental Material published in *EHP* articles may not conform to 508 standards due to the complexity of the information being presented. If you need assistance accessing journal content, please contact ehponline@niehs.nih.gov. Our staff will work with you to assess and meet your accessibility needs within 3 working days.

collection during one or two days (Fisher et al. 2015). In the context of pregnancy, estimations of the within-subject variability of phenols other than bisphenol A are lacking.

Accurate description of the variability of phenol urinary concentrations during pregnancy is crucial for adopting biospecimens sampling strategies that limits exposure misclassification in etiological studies. Our aim was consequently to characterize the within-day, between-day (within a week), and between-week variability of urinary concentrations of 10 phenols in pregnant women.

## Methods

### *Study Participants*

This study relied on a subgroup of the feasibility study conducted between July 2012 and July 2013 in the planning of the SEPAGES cohort (Suivi de l'Exposition à la Pollution Atmosphérique durant la Grossesse et Effets sur la Santé [Assessment of Air Pollution Exposure during Pregnancy and Effects on Health]). In this feasibility study, 40 women with a singleton pregnancy and living in the Grenoble urban area (France) were recruited from private obstetrical practices, before 17 gestational weeks (calculated from the date of the last menstrual period). The exclusion criteria included inability to write or speak French, being <18 y of age, planning to give birth outside of one of the four maternity hospitals of the Grenoble urban area, and not being enrolled in the French social security system. All participating women and their partners provided written informed consent for themselves and their offspring for biological measurements and data collection (Ouidir et al. 2015). The SEPAGES feasibility cohort was approved by the appropriate ethical committees [CPP (Comité de Protection des Personnes Sud-Est); CNIL (Commission Nationale de l'Informatique et des Libertés); CCTIRS (Comité Consultatif sur le Traitement de l'Information en matière de Recherche dans le domaine de la Santé); ANSM (Agence Nationale de sécurité du Médicament et des produits de santé)]. The involvement of the Centers for Disease Control and Prevention (CDC) laboratory did not constitute engagement in human subject research.

### *Study Design and Urine Collection*

The urine collection protocol is described in Figure 1. Urine collection took place during 7 consecutive days at three periods of pregnancy (first collection week, median: 13 gestational weeks, minimum–maximum [min–max]: 10–18 gestational weeks; second collection week, median: 23, min–max: 21–26; and third collection week, median: 32, min–max: 29–33; Table 1). Thirty of the 40 women participating in the SEPAGES feasibility study were asked to collect ~60 mL of each urine void, and to report in daily diaries micturition time for collected and missing voids. The remaining 10 participating women were asked to collect ~60 mL of only three urine voids per day, and were therefore not considered in the present study. The women collected the urine in polypropylene containers and stored it in a refrigerator (4°C [39°F]) in their home. When they were not at home, collected urine was stored in a cooler with ice packs. Specimens were retrieved two or three times a week by the study staff and brought in coolers to the Inserm research center [Institute for Advanced Biosciences (IAB), Grenoble, France]. Each sample was aliquoted into 2-mL polypropylene cryovials (up to five vials per sample) and frozen at −80°C (−112°F) until undergoing the pooling procedure or shipping for analysis. Because of cost constraints, we quantified only phenol biomarkers in the subgroup of eight women with the smallest rate of missed voids. Among these women, two had managed to collect a sample of each of their urine voids (no missing void, subgroup 1), while the other six women collected more than 95% of their weekly urine voids (subgroup 2).

### *Pooling Procedure*

We thawed at 4°C (39°F) and vortexed aliquots in polypropylene containers and pooled them according to the protocol detailed in Figure 1. For each woman and each study day, we took equal volumes of urine from all samples of the day and combined them within-woman, leading to seven within-subject daily pools for a 1-wk period (days 1 to 7). For each subject, we then prepared three weekly pools by combining an equal volume of the seven daily pools from each collection week (weeks 1, 2, and 3).

Immediately after preparation, pooled samples were placed in 2-mL polypropylene cryovials and frozen at −80°C (−112°F). The pools and all aliquots from spot samples to be analyzed were kept frozen until shipment on dry ice to the CDC laboratory in Atlanta (Georgia, USA). At the CDC laboratory, all urine samples were stored at or below −70°C (−94°F) until analysis.

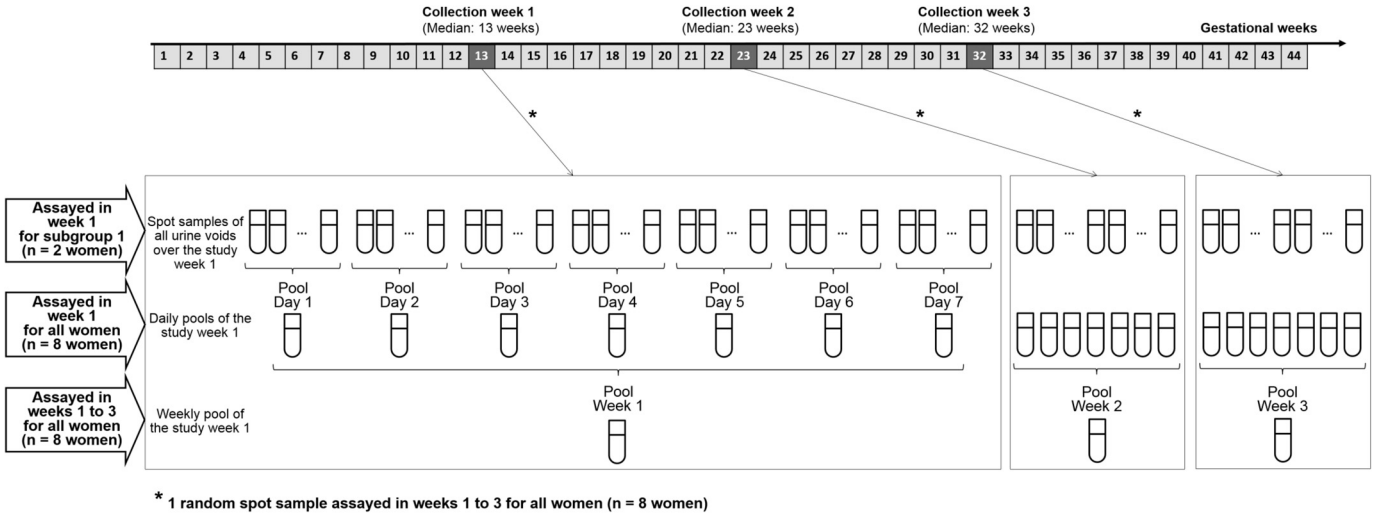


**Figure 1.** Urine collection, pooling procedure, and biomarker assays in all the study population ($N = 8$ women) and in the nested subgroup 1 ($N = 2$ women).

**Table 1.** Characteristics of the population (40 pregnant women from SEPAGES feasibility study, out of which 8 participated in the current study).

| Characteristic | Included ($n=8$) $N$ (%) or mean ± SD | Excluded ($n=32$) $N$ (%) or mean ± SD | $p$-Value[a] |
|---|---|---|---|
| Maternal age at enrollment (y) | 29.6 ± 3.8 | 30.7 ± 3.6 | 0.34 |
| Civil status | | | 0.25 |
| Married | 5 (63) | 12 (37) | |
| Cohabitating | 3 (37) | 20 (63) | |
| Maternal education | | | 1.00 |
| High school or less | 0 | 2 (6) | |
| ≤3 y of college | 4 (50) | 14 (44) | |
| >3 y of college | 4 (50) | 16 (50) | |
| Smoking history during pregnancy | | | 1.00 |
| Yes | 1 (12) | 6 (19) | |
| No | 7 (88) | 21 (66) | |
| Missing | 0 | 5 (15) | |
| Parity | | | 1.00 |
| Primiparous | 5 (63) | 19 (59) | |
| Multiparous | 3 (37) | 13 (41) | |
| Gestational age (wk) | | | |
| Week 1 of urine collection | 15.0 ± 1.9 | 12.9 ± 1.5 | 0.01 |
| Week 2 of urine collection | 24.0 ± 1.6 | 23.3 ± 1.4 | 0.21 |
| Week 3 of urine collection | 32.3 ± 0.7 | 31.8 ± 0.9 | 0.16 |
| Time between two weeks of urine collection (wk) | | | |
| Week 1–Week 2 | 9.0 ± 1.9 | 10.5 ± 1.8 | 0.06 |
| Week 2–Week 3 | 8.3 ± 1.2 | 8.5 ± 1.7 | 0.88 |

Note: SD, standard deviation; SEPAGES, Suivi de l'Exposition à la Pollution Atmosphérique durant la Grossesse et Effets sur la Santé (Assessment of Air Pollution Exposure during Pregnancy and Effects on Health).
[a]$p$-Values of chi-square or Mann-Whitney $U$-tests comparing the characteristics of included and non-included women.

### *Phenols, Creatinine, and Specific Gravity Measurements*

The total urinary concentrations of 2,4- and 2,5-dichlorophenols, benzophenone-3, bisphenol A, bisphenol S, triclosan, and butyl, methyl, ethyl, and propylparabens were quantified at the CDC using a modification of an online solid-phase extraction high-performance liquid chromatography-isotope dilution-tandem mass spectrometry method (Zhou et al. 2014). Limits of detection (LODs) are listed in Table 2. The coefficients of variation of quality control measurements ranged, according to compounds and concentration ranges, between 3.4% and 14.7%. It was higher for methylparaben, bisphenol A, benzophenone-3, and triclosan, for which the range was 5.8–14.7% and lower for the other compounds (3.4–6.4%). Moreover, the team in charge of urine collection added eight replicates to the samples assayed for phenols, in a way that was blinded to the lab. Correlation coefficients between biomarker concentrations in samples and their replicates ranged from 0.95 (bisphenols A and S) to 1.00 (2,4-dichlorophenol, ethyl, propyl and butylparabens, and triclosan). Two urine dilution markers were also quantified in the same samples: urinary creatinine measured at the CDC using a Roche/Hitachi MODULAR ANALYTICS Urine Work Area (SWA) P (photometric analysis) module (Roche Diagnostics); and urinary specific gravity, measured at room temperature using a handheld Atago PAL 10-S refractometer (Atago) at the Inserm Grenoble laboratory.

We analyzed a total of 216 samples (136 spot samples, $8 \times 7 = 56$ daily pools, and $8 \times 3 = 24$ weekly pools; Figure 1) for phenol biomarkers, creatinine, and specific gravity. For all women, we analyzed the seven daily pools of the first study week (56 daily pools) and the three weekly pools (24 weekly pools). Additionally, for the two women in subgroup 1, we also analyzed all spot samples of week 1 (total, 114 samples)

**Table 2.** Within-day variability – descriptive statistics of the non-transformed biomarker concentrations (μg/L) for the unpooled spot samples from subgroup 1 (2 women, $n=114$ spot samples collected over the first week of collection), and ICCs based on $\log_{10}$-transformed phenol biomarker concentrations, creatinine concentration, and specific gravity. Values were not standardized for creatinine or specific gravity.

| Biomarker | LOD (μg/L) | Results above the LOD, $N$ (%) | Percentiles 5th | 25th | 50th | 75th | 95th | Within-day ICC (95% CI)[a] |
|---|---|---|---|---|---|---|---|---|
| Phenols (μg/L)[b] | | | | | | | | |
| 2,4-Dichlorophenol | 0.1 | 113 (99) | 0.1 | 0.2 | 0.3 | 0.3 | 0.5 | 0.12 (0.00, 0.28) |
| 2,5-Dichlorophenol | 0.1 | 111 (97) | 0.1 | 0.2 | 0.4 | 0.7 | 1.2 | 0.11 (0.00, 0.27) |
| Butyl paraben | 0.1 | 113 (99) | 0.1 | 1.1 | 4.2 | 21.7 | 92.7 | 0.10 (0.00, 0.25) |
| Ethyl paraben | 1.0 | 105 (92) | <LOD | 2.4 | 10.7 | 43.1 | 126.2 | 0.03 (0.00, 0.15) |
| Methyl paraben | 1.0 | 114 (100) | 37 | 90.4 | 217.3 | 1329.4 | 5000.0 | 0.27 (0.05, 0.49) |
| Propyl paraben | 0.1 | 114 (100) | 0.8 | 4.8 | 36.7 | 139.0 | 895.1 | 0.28 (0.05, 0.50) |
| Benzophenone-3 | 0.2 | 31 (27) | <LOD | <LOD | <LOD | 1.4 | 10.7 | 0.26 (0.04, 0.48) |
| Bisphenol A | 0.1 | 113 (99) | 0.3 | 0.9 | 1.6 | 2.4 | 7.5 | 0.21 (0.01, 0.41) |
| Bisphenol S | 0.1 | 112 (98) | 0.1 | 0.2 | 0.3 | 0.5 | 1.3 | 0.50 (0.26, 0.73) |
| Triclosan | 1.0 | 90 (79) | <LOD | 1.0 | 2.1 | 3.3 | 4.8 | 0.30 (0.08, 0.53) |
| Urine dilution markers | | | | | | | | |
| Creatinine (mg/dL) | NA | 114 (100) | 37.2 | 72.5 | 102.5 | 139.2 | 223.1 | 0.10 (0.00, 0.26) |
| Specific gravity | NA | 114 (100) | 1.009 | 1.014 | 1.018 | 1.021 | 1.027 | 0.03 (0.00, 0.15) |

Note: CI, confidence interval; ICC, intraclass correlation coefficient; LOD, limit of detection; NA, not applicable.
[a]ICCs were estimated from ANOVA with a random effect on day (14 d) and within-woman mean-centering of the log-transformed phenols concentrations.
[b]Concentrations below the LOD were replaced by instrumental reading values. For each phenol biomarker, instrumental reading values equal to 0 were replaced by the lowest non-zero instrumental reading value divided by the square root of 2.

and 2 spot samples randomly selected among those from the two other collection weeks (1 random spot sample in each collection week). Finally for each of the six participants in subgroup 2, we analyzed 3 spot samples randomly selected among all the samples from the three collection weeks (one in each week of collection; Figure 1), so that, for all eight women, we could assay 1 random spot sample for each of the three collection weeks.

### *Statistical Analyses*

Concentrations below the LOD were replaced by instrumental readings, or by the compound-specific lowest non-zero instrumental reading divided by the square root of 2 when the instrumental reading was zero. We $\log_{10}$-transformed the urinary concentrations of phenol biomarkers to achieve approximate normality in the distributions. Correlations of biomarker concentrations between types of sample (unpooled samples, daily and weekly pools) for a given biomarker and between biomarker concentrations for a given type of sample were calculated using Spearman correlation coefficients.

Our assessment of variability relied on intraclass correlation coefficients (ICCs) estimated using one-way random-effect ANOVA models. The approach was identical for all ICCs estimations. ICCs close to zero indicate poor reproducibility of a concentration within the considered period, whereas values close to one indicate high reproducibility. In the case of a negative ICC estimate (which can happen with ANOVA models; Wang et al. 1992), we considered the ICC not to be computable and only reported the 95% confidence interval (CI), truncating its lower bound to zero.

To characterize the within-day variability, we defined ICC as the ratio of the between-day variance to the total variance (sum of within- and between-day variances). We relied on women of subgroup 1 (a total of 114 samples, collected during the first collection week for the two women). The woman- and compound-specific weekly mean was subtracted from the spot concentrations to correct for the between-subject variability before estimating ICCs representing the within-day variability.

We assessed the within-week (between-day) and between-week variability with ICCs calculated as the ratio of the between-subject variance to the total variance (sum of within- and between-subject variance). For the between-day variability, we ran models based on all 56 daily pools of the study week 1 (eight women, each with seven daily pools). For the between-week variability, we used all 24 weekly pools (eight women, each with three weekly pools).

To allow comparisons with previous studies that relied on two to three spot samples collected during pregnancy, we additionally estimated ICCs based on the three random spot samples collected during pregnancy (each sample being randomly selected in each collection week for the eight women). To assess the robustness of the findings to the statistical methods, we also computed ICCs using random intercept linear mixed models (maximum likelihood estimates) instead of ANOVA models. To assess the potential impact of urinary dilution on variance estimates for phenol biomarkers, the ANOVA analyses were repeated using phenol biomarker concentrations corrected for creatinine (ratio of the phenol biomarker concentrations to the creatinine concentration in the same sample) and for specific gravity using a formula previously described (Philippat et al. 2013), or by including creatinine concentration or specific gravity as a covariate in the random intercept linear mixed models. Data were analyzed using STATA (version 12.1; Stata Corporation).

## Results

### *Study Population and Samples*

At enrollment, women were 29.6 y of age on average [standard deviation (SD): 3.8]; most of them were primiparous (63%) and did not smoke during pregnancy (88%) and all of them had college education (Table 1). Women collected from 3 to 15 urine samples per day (total from 132 to 240 samples per woman). Detection frequencies were generally between 79% (triclosan) and 100% (methyl and propylparabens), except for benzophenone-3, which was only detected in 35% of the samples. The highest (between-compound) coefficients of correlation were observed between structurally similar compounds (e.g., between the two dichlorophenols and between the four parabens), and between creatinine and specific gravity (see Table S1), regardless of the type of sample (unpooled, daily, or weekly pools). Correlation between creatinine and urinary concentrations of phenol biomarkers ranged from −0.06 (with propylparaben) to 0.83 (2,4-dichlorophenol) in spot samples of subgroup 1, from −0.28 (triclosan) to 0.61 (bisphenol A) in daily pools of all studied women, and from −0.15 (triclosan) to 0.56 (2,4-dichlorophenol) in weekly pools.

### *Within-Day Variability*

For the spot samples, urinary concentrations of most phenols and of creatinine varied within-woman by several orders of magnitude throughout the first collection week and within a day (Figure 2). For all biomarkers including creatinine and specific gravity, ICCs were low to moderate (Table 2), with the highest ICC observed for bisphenol S (0.50; 95% CI: 0.26, 0.73), and the lowest for ethylparaben (0.03; 95% CI: 0.00, 0.15) and specific gravity (0.03; 95% CI: 0.00, 0.15). Creatinine or specific gravity standardization, respectively, slightly increased (by 0.02 to 0.07) ICCs for 5 and 9 compounds out of 10 (see Table S2). Using linear mixed models instead of ANOVA methods to estimate ICCs (see Tables S3 and S4) led to similar results.

### *Between-Day (Within a Week) Variability*

The remaining analyses are based on the whole group of eight women. Detection frequencies were quite similar in daily pools (median over all compounds, 100%) compared with unpooled spot samples (median, 99%) for most of the compounds, except for butylparaben (80% detection in daily pools, compared with 99% in spot samples), ethylparaben (70% vs. 92%), and benzophenone-3 (50% vs. 27%; Table 3). ICCs based on daily pools were >0.80 high for most compounds, except for benzophenone-3 (0.73; 95% CI: 0.50, 0.96), bisphenol A (0.60; 95% CI: 0.30, 0.89) and bisphenol S (0.14; 95% CI: 0.00, 0.39; Table 3 and Figure 3). The highest ICCs were observed for 2,4-dichlorophenol (0.91; 95% CI: 0.82, 1.00) and 2,5-dichlorophenol (0.98; 95% CI: 0.95, 1.00). ICCs for creatinine and specific gravity were greater in daily pools (0.60 and 0.61, respectively) compared with spot samples (0.10 and 0.03, respectively). For five compounds, correction for creatinine slightly increased (by 0.01 to 0.05) ICCs, which stayed within ranges of the uncorrected confidence intervals (see Table S5). We observed similar results using a linear mixed model instead of ANOVA methods to estimate ICCs (see Tables S3 and S6).

### *Between-Week Variability Based on Weekly Pools*

On average, 9.0 wk (SD: 1.9) elapsed between collection weeks 1 and 2, and 8.3 wk (SD: 1.2) between collection weeks 2 and 3 (range, 5.6–12.0 wk). For 7 of the 10 phenol biomarkers, detection frequencies in weekly pools were similar to those in daily pools (>79%; Table 4), whereas they were somewhat higher in weekly pools for butyl- and ethylparabens (88% and 79%,

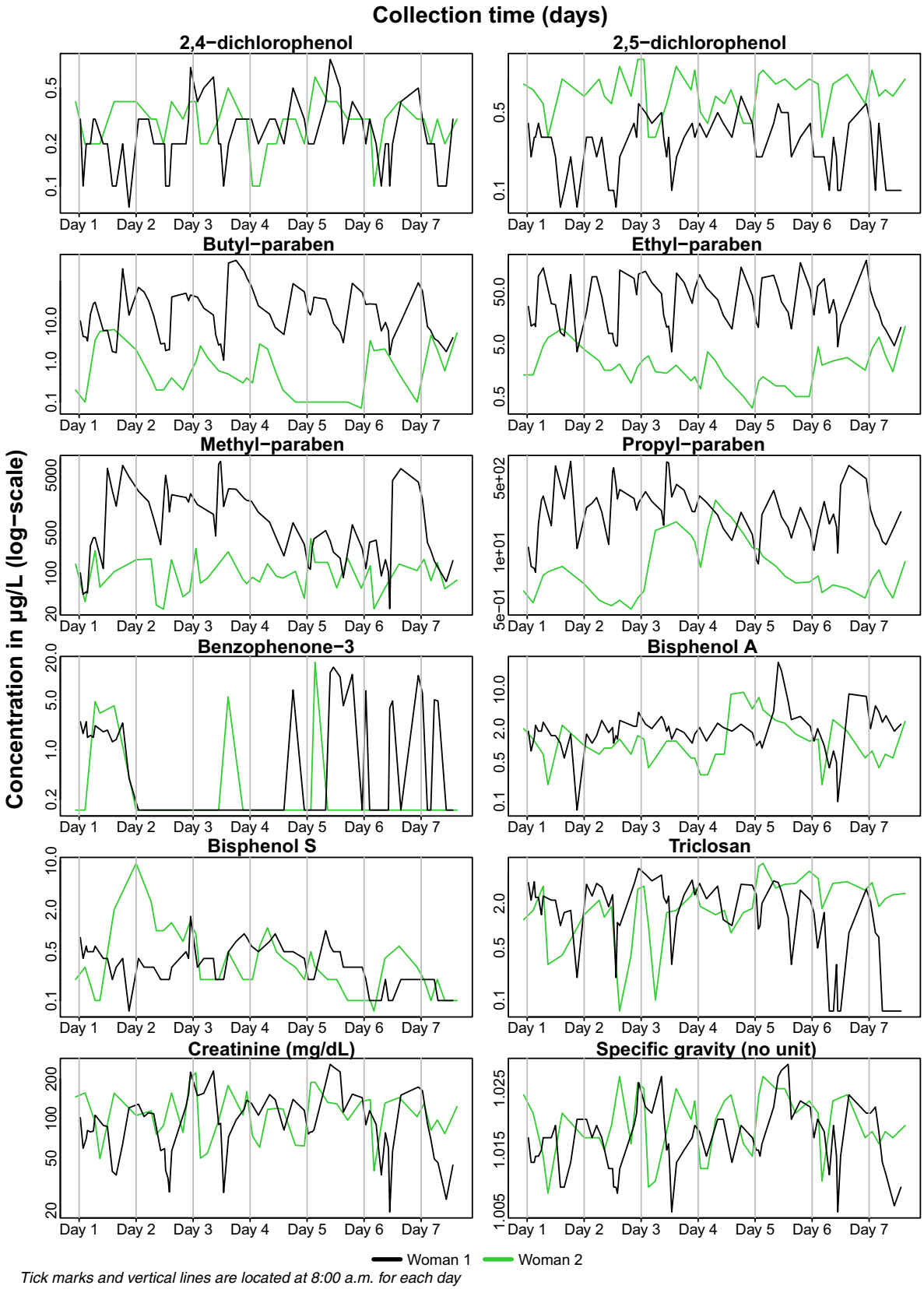


**Figure 2.** Within-day and within-week variability – urinary concentrations of 10 phenols (µg/L), creatinine concentration (mg/dL), and specific gravity in $\log_{10}$-scale in the unpooled spot samples from subgroup 1 (2 women, $n = 114$ spot samples collected over the first week of collection). Note that to facilitate visualization, each biomarker is displayed on a specific scale.

**Table 3.** Between-day variability – descriptive statistics of the non-transformed biomarker concentrations (μg/L) for the within-woman daily pooled samples from subgroup 1 and subgroup 2 (8 women, $n = 56$ daily pools, one daily pool for each day of the first week of collection), and ICCs based on $\log_{10}$-transformed phenol biomarker concentrations, creatinine concentration, and specific gravity. Values were not standardized for creatinine or specific gravity.

| Biomarker | LOD (μg/L) | Results above the LOD, *N* (%) | Percentiles | | | | | Between-day ICC (95% CI)[a] |
|---|---|---|---|---|---|---|---|---|
| | | | 5th | 25th | 50th | 75th | 95th | |
| Phenols (μg/L)[b] | | | | | | | | |
| 2,4-Dichlorophenol | 0.1 | 56 (100) | 0.1 | 0.2 | 0.3 | 0.4 | 6.0 | 0.91 (0.82, 1.00) |
| 2,5-Dichlorophenol | 0.1 | 56 (100) | 0.2 | 0.3 | 0.8 | 3.6 | 269.6 | 0.98 (0.95, 1.00) |
| Butyl paraben | 0.1 | 45 (80) | 0.1 | 0.1 | 0.2 | 1.1 | 35.2 | 0.80 (0.61, 0.99) |
| Ethyl paraben | 1.0 | 39 (70) | <LOD | <LOD | 2.8 | 12.9 | 45.5 | 0.85 (0.70, 1.00) |
| Methyl paraben | 1.0 | 56 (100) | 2.4 | 9.2 | 33.0 | 89.7 | 1350.8 | 0.84 (0.69, 1.00) |
| Propyl paraben | 0.1 | 56 (100) | 0.1 | 0.4 | 2.2 | 39.9 | 161.8 | 0.90 (0.80, 1.00) |
| Benzophenone-3 | 0.2 | 28 (50) | <LOD | <LOD | 0.5 | 3.1 | 33.8 | 0.73 (0.50, 0.96) |
| Bisphenol A | 0.1 | 56 (100) | 0.4 | 1.0 | 1.9 | 3.5 | 8.2 | 0.60 (0.30, 0.89) |
| Bisphenol S | 0.1 | 56 (100) | 0.1 | 0.2 | 0.2 | 0.4 | 2.8 | 0.14 (0.00, 0.39) |
| Triclosan | 1.0 | 45 (80) | <LOD | 1.1 | 2.1 | 3.0 | 63.7 | 0.89 (0.78, 1.00) |
| Urine dilution markers | | | | | | | | |
| Creatinine (mg/dL) | NA | 56 (100) | 30.6 | 60.4 | 77.9 | 102.2 | 146.8 | 0.60 (0.30, 0.89) |
| Specific gravity | NA | 56 (100) | 1.009 | 1.013 | 1.016 | 1.02 | 1.029 | 0.61 (0.32, 0.90) |

Note: CI, confidence interval; ICC, intraclass correlation coefficient; LOD, limit of detection; NA, not applicable.
[a]ICCs were estimated from ANOVA with a random effect on woman.
[b]Concentrations below the LOD were replaced by instrumental reading values. For each phenol biomarker, instrumental reading values equal to 0 were replaced by the lowest non-zero instrumental reading value divided by the square root of 2.

compared with 80% and 70%, respectively, in daily pools), and lower for benzophenone-3 (38% vs. 50% in daily pools). Between two study weeks, concentrations of almost all phenols varied by several orders of magnitude for some women (Figure 4). ICCs for 2,5-dichlorophenol; butyl-, methyl-, and propylparabens; and creatinine were >0.8, while they were <0.6 for the other biomarkers (Table 4, Figure 5), and lowest for bisphenol S (0.26; 95% CI: 0.00, 0.73). ICC for triclosan could not be computed, but was probably in the low range (95% CI: 0.00, 0.44). ICCs slightly decreased (by 0.01 to 0.2) for most compounds when using creatinine- or specific gravity–corrected concentrations, but stayed within ranges of the uncorrected confidence intervals (see Table S5). Adjusting for creatinine or specific gravity did not change the results (see Tables S3 and S6).

### *Between-Week Variability Based on Three Random Spot Samples*

Detection frequencies of biomarker concentrations in the three random spot samples collected in each woman were >79%, except for ethylparaben (70%) and triclosan (54%). ICCs computed using these random spots samples were lower than 0.6, except for 2,5-dichlorophenol (0.85; 95% CI: 0.69, 1.00), methylparaben (0.85; 95% CI: 0.68, 1.00), and propylparaben (0.70; 0.40, 1.00) (Table 5). As for weekly pools, ICCs slightly decreased (by 0.01 to 0.19) when using creatinine- or specific gravity–corrected concentrations (see Table S5). Adjustment for creatinine or specific gravity did not change the results (see Tables S3 and S6).

## Discussion

To our knowledge, this study is the first to evaluate the within-day, between-day, and between-week variability of 10 phenol biomarkers as well as of creatinine and specific gravity in pregnant women. Most compounds showed very high variability over the course of a day (ICCs generally <0.3), while the between-day variability of the daily averages over the course of a week was much lower. This pattern was opposite for bisphenol S, which had a stronger between-day than within-day variability. The variability of the weekly averages considered several weeks apart exhibited more contrasted patterns across compounds, with low between-week variability for some compounds (2,5-dichlorophenol, butylparaben, propylparaben, methylparaben) and a high variability for others (ethylparaben, bisphenol S, triclosan). Urinary dilution or creatinine levels did not explain much of the observed within-subject variability in phenol biomarkers.

### *Strengths and Limitations*

A key strength of our study is the reliance on pregnant women who agreed to collect samples from each micturition over three weeks. This study considered a large number of phenols, including bisphenol S and others for which the literature is rather sparse. Also, contrary to previous studies, our design allowed characterization of the temporal variability of phenol biomarkers over several time windows during pregnancy, and in particular within the day and between the days of a week. From eight blinded samples analyzed in duplicate, we observed very high (>0.95) correlation between two analyses of the same urine sample for all biomarkers, making it very unlikely that the reported ICCs were strongly influenced by analytical error. A limitation relates to the fact that our estimate of the within-day variability relied on samples collected by only two women, contrary to the estimates of the between-day and between-week variability, which relied on eight women. Volumes of urine voids were not collected, preventing us from calculating the excretion rate. Caution is required in interpreting estimates for benzophenone-3, which was the compound with the lowest detection frequency in all samples (27% to 54%), and in interpreting the results related to the analyses of the weekly pools, given the large confidence intervals. Even though we restricted our study population to women with only a few missed voids (<5%), pools were created using all available urine samples and these missing voids may be a source of error. Given that the present study was restricted to a specific population and specific chemicals, generalization of our results to other populations or other compounds should be considered with great caution.

### *Study Population*

Our study relied on a small population of women who agreed to collect repeated urine samples for several weeks. It was not meant to be representative of the general population or of all pregnant women from France or even the Grenoble area. Among those approached to participate, women with a high education level or interest in environmental or health issues were more

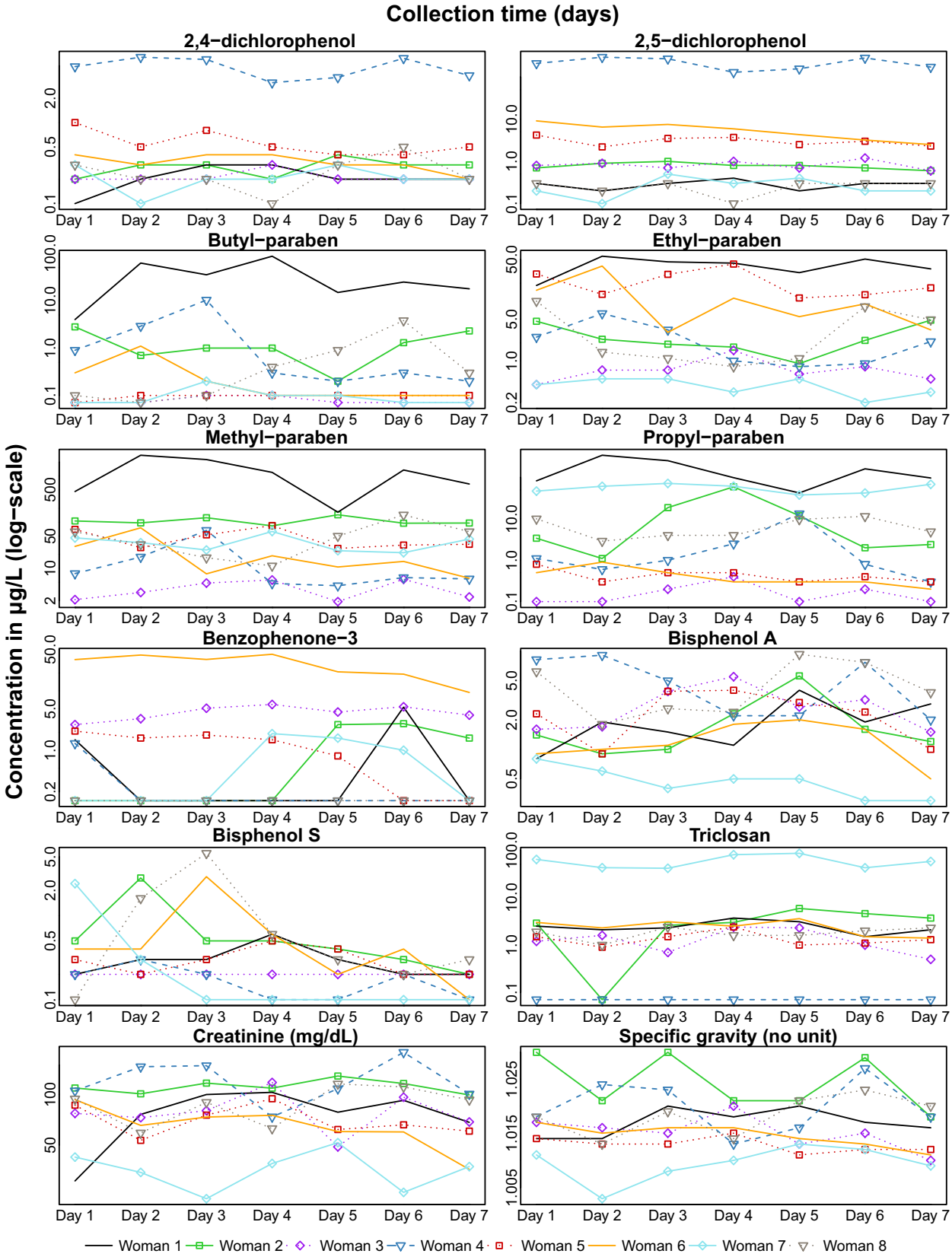


**Figure 3.** Between-day (within a week) variability of pooled daily samples – urinary concentrations of 10 phenols (µg/L), creatinine concentration (mg/dL), and specific gravity in $\log_{10}$-scale in the within-woman daily pooled samples from subgroup 1 and subgroup 2 (8 women, $n = 56$ daily pools, one daily pool for each day of the first week of collection). Note that to facilitate visualization, each biomarker is displayed on a specific scale.

**Table 4.** Between-week variability – descriptive statistics of the non-transformed biomarker concentrations (μg/L) for the within-woman weekly pooled samples from subgroup 1 and subgroup 2 (8 women, $n = 24$ weekly pools, one weekly pool for each of the three weeks of collection), and ICCs based on $\log_{10}$-transformed phenol biomarker concentrations, creatinine concentration, and specific gravity. Values were not standardized for creatinine or specific gravity.

| Biomarker | LOD (μg/L) | Results above the LOD, $N$ (%) | Percentiles | | | | | Between-week ICC (95% CI)[a] |
|---|---|---|---|---|---|---|---|---|
| | | | 5th | 25th | 50th | 75th | 95th | |
| Phenols (μg/L)[b] | | | | | | | | |
| 2,4-Dichlorophenol | 0.1 | 24 (100) | 0.1 | 0.2 | 0.3 | 0.6 | 2.2 | 0.65 (0.32, 0.99) |
| 2,5-Dichlorophenol | 0.1 | 24 (100) | 0.3 | 0.4 | 0.6 | 2.2 | 73.0 | 0.93 (0.86, 1.00) |
| Butyl paraben | 0.1 | 21 (88) | 0.1 | 0.1 | 0.3 | 0.7 | 24.0 | 0.84 (0.67, 1.00) |
| Ethyl paraben | 1.0 | 19 (79) | <LOD | 1.1 | 11.4 | 19.5 | 56.5 | 0.33 (0.00, 0.79) |
| Methyl paraben | 1.0 | 24 (100) | 3.9 | 23.7 | 44.4 | 99.5 | 1673.7 | 0.81 (0.60, 1.00) |
| Propyl paraben | 0.1 | 24 (100) | 0.2 | 0.4 | 4.8 | 44.4 | 173.6 | 0.86 (0.70, 1.00) |
| Benzophenone-3 | 0.2 | 9 (38) | <LOD | <LOD | <LOD | 4.3 | 28.5 | 0.60 (0.23, 0.97) |
| Bisphenol A | 0.1 | 24 (100) | 0.5 | 1.3 | 1.9 | 3.3 | 5.7 | 0.59 (0.22, 0.97) |
| Bisphenol S | 0.1 | 24 (100) | 0.1 | 0.2 | 0.3 | 0.7 | 14.4 | 0.26 (0.00, 0.73) |
| Triclosan | 1.0 | 19 (79) | <LOD | 1.3 | 2.5 | 7.1 | 83.7 | (0.00, 0.44) |
| Urine dilution markers | | | | | | | | |
| Creatinine (mg/dL) | NA | 24 (100) | 49.6 | 66.0 | 84.1 | 111.8 | 142.3 | 0.83 (0.64, 1.00) |
| Specific gravity | NA | 24 (100) | 1.009 | 1.012 | 1.016 | 1.021 | 1.023 | 0.49 (0.07, 0.91) |

Note: CI, confidence interval; ICC, intraclass correlation coefficient; LOD, limit of detection; NA, not applicable.
[a]ICCs were estimated from ANOVA with a random effect on woman. For triclosan, we give only the confidence interval truncated to zero due to negative estimate of ICC.
[b]Concentrations below the LOD were replaced by instrumental reading values. For each phenol biomarker, instrumental reading values equal to 0 were replaced by the lowest nonzero instrumental reading value divided by the square root of 2.

likely to participate. Consequently, the behaviors (use of personal care products, diet, etc.) of our population are unlikely to represent those of all pregnant women. One might anticipate that this possibly led to an underestimation of the between-woman variability in urinary concentrations for some of the considered compounds, although overrepresenting women using few personal care products (as may be the case for some highly educated women aware of the health concerns regarding the use of health care products during pregnancy) or with a diet low in industrial phenols may also have led to underestimating the within-subject variability. Other factors such as physical activity, which might influence the toxicokinetics of xenobiotics (Persky et al. 2003), may also have differed in our population. Most phenol urinary concentrations were lower than those reported in previous cohorts of pregnant women (Guidry et al. 2015; Meeker et al. 2013; Philippat et al. 2013; Smith et al. 2012) and in 1,230 U.S. nonpregnant women in 2011–2012 (CDC 2015). In addition to differences in behaviors, composition of consumers' products, regulation of chemicals in each country, and analytical methods across laboratories performing the assays may also have differed across studies.

### *Variability over the Course of Pregnancy*

Our analyses based on three random spot samples were meant to describe the ability of a simple sampling approach to capture the whole pregnancy exposure, and to allow comparison with previous studies, which relied on up to three spot samples per participant to assess biomarker concentration variability during pregnancy (Bertelsen et al. 2014; Braun et al. 2011, 2012; Guidry et al. 2015; Jusko et al. 2014; Meeker et al. 2013; Philippat et al. 2013; Smith et al. 2012; Stacy et al. 2017). None of these studies investigated bisphenol S. The moderate ICCs (between 0.4 and 0.5) observed for 2,4-dichlorophenol and butyl- and ethylparaben urinary concentrations were consistent with previous reports (Guidry et al. 2015; Meeker et al. 2013; Philippat et al. 2013; Smith et al. 2012), while we observed greater ICCs for 2,5-dichlorophenol and methylparaben (>0.8 compared with 0.4–0.5 in these previous studies). Also, considering their rather large confidence intervals, ICCs for bisphenol A (0.4; 95% CI: 0.0, 0.8) and propylparaben (0.7; 95% CI: 0.4, 1.0) were within the range of previous studies ($\sim$0.3 for bisphenol A and from 0.3 to 0.6 for propylparaben), although at the upper end of the range. In contrast, ICCs for benzophenone-3 (0.3; 95% CI: 0.0, 0.8) and triclosan (0.1; 95% CI: 0.0, 0.6) were at the lower end of the range of previously reported results, which were between 0.3 and 0.6 (Bertelsen et al. 2014; Meeker et al. 2013; Philippat et al. 2013; Stacy et al. 2017). Detection rates in our population were low (<55%) for these two compounds, that might have decreased ICCs compared with studies with higher detection rates, due to more homogeneity between women, and hence, a proportionally larger within-subject variability.

Our study is, to the best of our knowledge, the first to rely on within-subject weekly pools instead of random spot samples to describe the variability of select phenols during pregnancy. Compared with the results based on random spot samples from our study, ICCs based on weekly pools tended to be higher but the overall conclusion was similar than with the three samples approach: variability was contrasted between compounds; it was low for 7 of the 10 phenols (ICCs >0.59), and high (ICCs <0.4), for the other compounds (ethylparaben, bisphenol S, and triclosan).

### *Between-Day (within-Week) Variability*

For most biomarkers, the between-day variability of the urinary concentrations over a week was low (ICCs >0.7 for 8 of the 10 phenols, and ICC of 0.6 for bisphenol A). For bisphenol S, between-day variability was high (0.14; 95% CI: 0.0, 0.39). To the best of our knowledge, no study had relied on within-subject daily pools to investigate the variability of phenol urinary concentrations over several consecutive days. Two previous studies had relied on 24-h simulated urine concentrations (volume-weighted averages of all daily urine voids) in a nonpregnant population of eight males and females (who collected all their complete urine voids and recorded urine volumes) to characterize the between-day variability of bisphenol A over a week (Ye et al. 2011) and of several phenols over four consecutive days (Koch et al. 2014). We observed a somewhat lower variability of bisphenol A urinary concentration in daily pools (0.6; 95% CI: 0.30, 0.89) compared with these previous studies in nonpregnant subjects (ICCs between 0.12 and 0.28), whereas for ethyl-, methyl-, and propylparabens; triclosan; and benzophenone-3, ICCs (between 0.73 and 0.98) were in close agreement with those

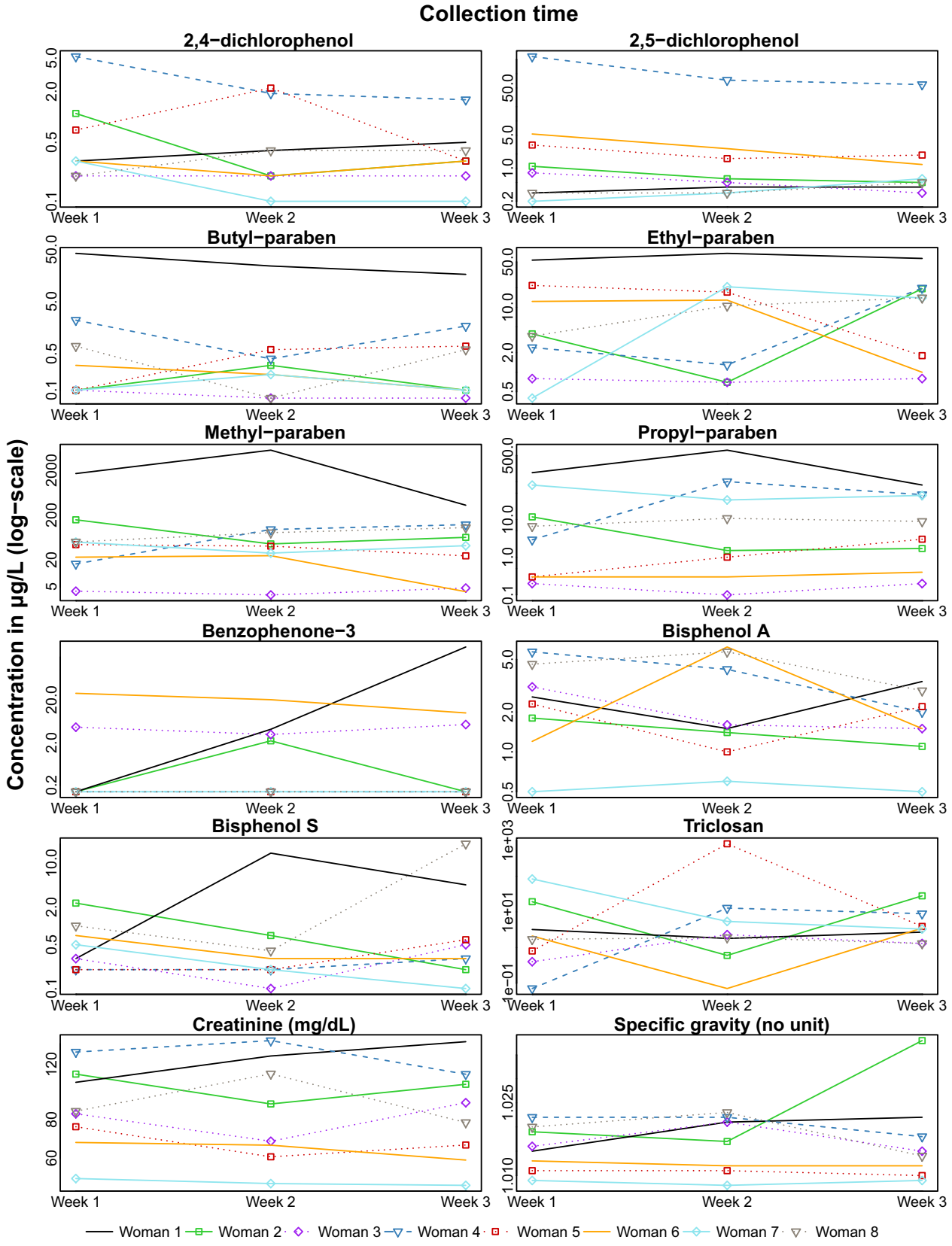


**Figure 4.** Between-week variability of weekly samples – urinary concentrations of 10 phenols (µg/L), creatinine concentration (mg/dL), and specific gravity in $\log_{10}$-scale in the within-woman weekly pooled samples from subgroup 1 and subgroup 2 (8 women, $n = 24$ weekly pools, one weekly pool for each of the three weeks of collection). Note that to facilitate visualization, each biomarker is displayed on a specific scale.

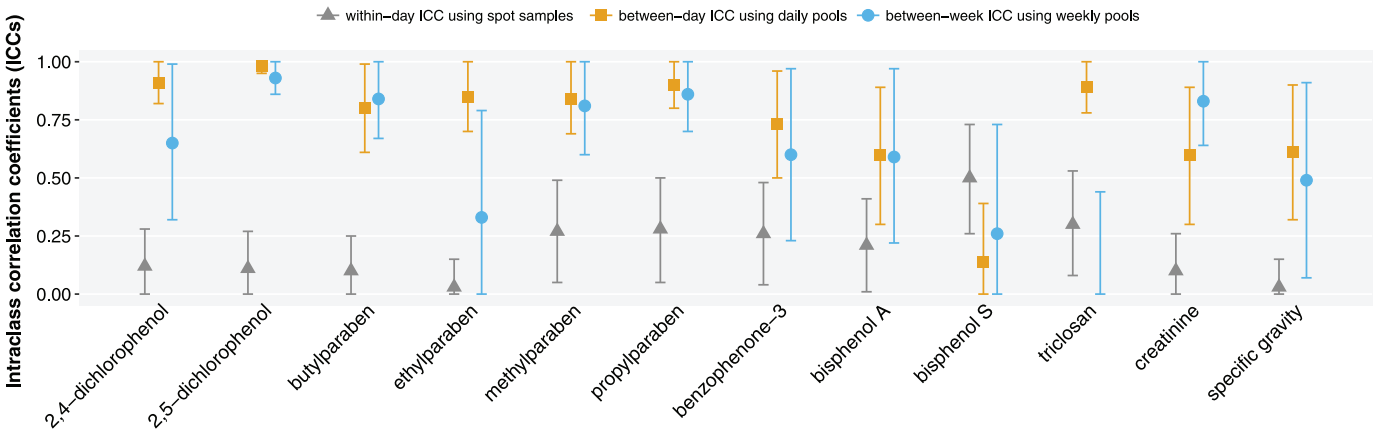


**Figure 5.** Intraclass correlation coefficients (error bars for 95% confidence intervals) for the within-day variability using the unpooled spot samples from subgroup 1 (2 women, $n = 114$ spot samples collected over the first week of collection, triangle markers), the between-day variability using the within-woman daily pooled samples from subgroup 1 and subgroup 2 (8 women, $n = 56$ daily pools, one daily pool for each day of the first week of collection, square markers), and the between-week variability in the within-woman weekly pooled samples from subgroup 1 and subgroup 2 (8 women, $n = 24$ weekly pools, one weekly pool for each of the three weeks of collection, circle markers). For triclosan, we give only the confidence interval truncated to zero due to the negative estimate of between-week ICC.

reported by Koch et al. (2014) (between 0.71 and 0.99). These findings suggest a good reproducibility of daily averages urinary concentrations over a week for most target phenols but the bisphenols.

### *Within-Day Variability*

Within-day variability was high for all phenol biomarkers (ICCs $\leq$0.50), showing that a random spot sample collected within a day does not accurately represent the daily average. For bisphenol A (0.21; 95% CI: 0.01, 0.41), this result is in line with findings from Ye et al. (2011), who measured urinary bisphenol A concentration in all spot urine samples collected in a 1-wk period from eight nonpregnant participants (ICCs, 0.12–0.21 for the within-day variability). This is also in agreement with a reported low reproducibility of bisphenol A concentrations in urine samples from 1 day (ICCs between 0.31 and 0.33) in a study in pregnant women who provided all their urine voids during 1 or 2 d (66 women) as well as spot samples at different time points during and after pregnancy (Fisher et al. 2015). The high within-day variations in phenol urinary concentrations are probably related to the (expectedly) very low half-life of phenols in pregnant women, as can be deduced from studies in nonpregnant human adults (Janjua et al. 2007; Sandborgh-Englund et al. 2006; Völkel et al. 2002), and to exposure being episodic, with the main suspected exposure sources being diet (for bisphenols in food containers, for parabens used as preservatives in some industrial food preparation) and personal care products (for parabens, triclosan, benzophenone-3). The fact that the within-day variability was higher than the between-day variability for most compounds could be due to the behaviors driving exposure tending to be similar from one day to another.

### *Urine Dilution and Within-Subject Pooling*

Creatinine and specific gravity are commonly used to adjust for urine dilution in studies relying on urinary biomarkers (Barr et al. 2005; Boeniger et al. 1993). The two markers were strongly correlated (coefficient of correlation, 0.86–0.92) in spot samples and daily and weekly pools. As previously reported for creatinine,

**Table 5.** Alternative estimate of between-week variability based on three random spot samples – descriptive statistics of the non-transformed biomarker concentrations (μg/L) for the random spot samples from subgroup 1 and subgroup 2 (8 women, $n = 24$ random spot samples, one sample in each of the three weeks of collection), and ICCs based on $\log_{10}$-transformed phenol biomarker concentrations, creatinine concentration, and specific gravity. Values were not standardized for creatinine or specific gravity.

| Biomarker | LOD (μg/L) | Results above the LOD, $N$ (%) | Percentiles | | | | | Between-week ICC (95% CI)[a] |
|---|---|---|---|---|---|---|---|---|
| | | | 5th | 25th | 50th | 75th | 95th | |
| Phenols (μg/L)[b] | | | | | | | | |
| 2,4-Dichlorophenol | 0.1 | 24 (100) | 0.1 | 0.2 | 0.4 | 0.5 | 2.2 | 0.50 (0.08, 0.92) |
| 2,5-Dichlorophenol | 0.1 | 23 (96) | 0.1 | 0.3 | 0.8 | 1.4 | 70.3 | 0.85 (0.69, 1.00) |
| Butyl paraben | 0.1 | 19 (79) | 0.1 | 0.1 | 0.1 | 3.8 | 8.0 | 0.42 (0.00, 0.87) |
| Ethyl paraben | 1.0 | 16 (67) | <LOD | <LOD | 2.4 | 10.5 | 21.1 | 0.40 (0.00, 0.85) |
| Methyl paraben | 1.0 | 24 (100) | 2.3 | 5.0 | 15.4 | 106.3 | 276.7 | 0.85 (0.68, 1.00) |
| Propyl paraben | 0.1 | 24 (100) | 0.1 | 0.5 | 4.2 | 34.5 | 99.3 | 0.70 (0.40, 1.00) |
| Benzophenone-3 | 0.2 | 13 (54) | <LOD | <LOD | 0.8 | 4.5 | 47.4 | 0.28 (0.00, 0.75) |
| Bisphenol A | 0.1 | 23 (96) | 0.1 | 0.4 | 1.1 | 2.6 | 8.0 | 0.38 (0.00, 0.83) |
| Bisphenol S | 0.1 | 22 (92) | 0.1 | 0.1 | 0.2 | 0.4 | 12.3 | 0.33 (0.00, 0.80) |
| Triclosan | 1.0 | 13 (54) | <LOD | <LOD | 1.1 | 2.4 | 29.9 | 0.11 (0.00, 0.58) |
| Urine dilution markers | | | | | | | | |
| Creatinine (mg/dL) | NA | 24 (100) | 13.90 | 28.65 | 67.43 | 125.94 | 156.25 | 0.54 (0.14, 0.94) |
| Specific gravity | NA | 24 (100) | 1.005 | 1.007 | 1.016 | 1.021 | 1.029 | 0.69 (0.39, 1.00) |

Note: CI, confidence interval; ICC, intraclass correlation coefficient LOD, limit of detection; NA, not applicable.
[a]ICCs were estimated from ANOVA with a random effect on woman. For triclosan analysis, an extreme value was excluded ($n = 23$ samples).
[b]Concentrations below the LOD were replaced by instrumental reading values. For each phenol biomarker, instrumental reading values equal to 0 were replaced by the lowest non-zero instrumental reading value divided by the square root of 2.

creatinine concentration and specific gravity had high within-subject variations throughout a 1-wk period (Boeniger et al. 1993). In our study, correcting phenol concentrations by either creatinine or specific gravity did not greatly improve ICCs, suggesting that these parameters do not explain much of the biomarkers' variability. Because the total urine volumes were unknown, we created within-subject daily pools using an equal volume of each spot sample collected each day and results might have differed if pooling volumes had been based on specific gravity or creatinine concentration. Creatinine is a body waste product primarily excreted by glomerular filtration (Barr et al. 2005; Boeniger et al. 1993; Cheung and Lafayette 2013). Excretion profiles of phenols are not well characterized in humans and specifically in pregnant women. However, the low correlation between creatinine or specific gravity and most of our exposure biomarkers may reflect that urinary excretion processes for creatinine and these phenols might differ, and that the dilution of urine samples may not affect substantially biomarker concentrations. Also, in some areas of research, as a replacement for creatinine standardization, correction for urine dilution relies on parameters other than creatinine or specific gravity. For example, the Integral Quotient Normalization approach in metabolomics relies on adjustment by the median value of all biomarkers (Lindon et al. 2007).

### *Implications for Sampling Strategies in Etiological Studies*

We confirmed that studies aiming at characterizing the health effects of pregnancy exposure to compounds with a high within-subject variability such as most of those considered here should generally collect several biospecimens per subject to reduce exposure misclassification (Perrier et al. 2016). More importantly, we report for the first time period-specific ICCs for select phenols, which can be used to refine the urine sampling scheme in epidemiologic studies aiming at characterizing the health effects of such exposures. As shown in Table 4, for some compounds (e.g., 2,5-dichlorophenol, several parabens), if a good estimate of the exposure averaged over a specific pregnancy week is available (e.g., through collection of a sample of all urine voids over this week), then this can conveniently be used as an estimate of the average exposure over all pregnancy weeks. For triclosan and bisphenol S, for which the between-week ICCs are <0.3, assessing exposure during a small number of weeks is unlikely to provide a reasonable estimate of the whole pregnancy exposure average. For these compounds, focusing on a few specific weeks of pregnancy may be inefficient, and one may rather consider collecting random samples during pregnancy to estimate the pregnancy average. Relying on the simulation by Perrier et al. (2016) and on the ICCs estimated in the current study, for compounds such as triclosan and bisphenol S, two to three dozen urine samples would be required, while for compounds such as 2,4-dichlorophenol, with an ICC close to 0.6 for the pregnancy window, pooling approximately five urine samples would allow for strongly limiting bias in the dose–response function. Relying on a spot urine sample, is, under the assumption of classical-type error, likely to induce an attenuation bias by 30% (for propylparaben) to 50% or more in the dose–response function if a single spot sample is used in etiological studies (Perrier et al. 2016). It is only for 2,5-dichlorophenol and methylparaben, for which ICCs based on three random spot samples during pregnancy (Table 5) were equal to 0.85, that using one or two spot urine samples collected randomly during pregnancy may provide a reasonable estimate of the whole pregnancy exposure average. If one is interested in an exposure window of a length of a week (for example the week when a specific fetal organ starts developing or at the end of which some biological parameter is assessed in the mother), then for dichlorophenols, triclosan, and parabens, assessing exposure during a single day of the week should do the job, whereas for benzophenone-3 and bisphenols A and S, it is safer to assess exposure during several days of the week. If one is now interested in assessing exposures over a specific day, collecting one spot sample is likely not enough for all of the studied phenols, given that the within-day ICCs were all <0.5.

For biomonitoring (and not etiological studies), there is no issue related to bias in dose–response functions, and collecting a spot biospecimen might be a good option if the population is large enough. Collecting a random sample rather than the first morning void is likely to provide a much better estimate of the population average.

## Conclusion

Biospecimens sampling strategy for accurate exposure assessment is a key issue in epidemiological studies based on short half-lived chemicals such as phenols. Our findings confirm that exposure misclassification may be high when collecting a small number of random spot samples. Future etiological studies should adopt a carefully thought-out design for the biospecimen sampling instead of using the default option of a single biospecimen per subject. Our results suggest that collecting more than one biospecimen per day for preferably several days during pregnancy is likely to allow a reduction in exposure misclassification.

## Acknowledgments

We thank L. Borges, clinical research assistant, and the staff from Grenoble Center for Clinical Investigation (CIC): J.-L. Cracowski, E. Hodaj, D. Abry, A. Tournier, J. Quentin, and N. Gonnet. We are grateful for the support of M. Althuser, F. Camus-Chauvet, D. Marchal André, X. Morin, P. Rivoire, A. Royannais, C. Tomasella, T. Tomasella, P. Viossat, E. Volpi, S. Rey, P. Hoffmann, and clinicians from the Grenoble University Hospital in the recruitment of the study volunteers. We also acknowledge the technical assistance of P. Dwivedi, X. Zhou, J. Tao, T. Powell, and P. Olive (CDC, Atlanta, GA) in measuring the urinary concentrations of phenol biomarkers and creatinine. We would also like to give our sincere thanks to participants of the SEPAGES feasibility study.

This work was supported by the European Research Council [ERC consolidator grant 311765 E-DOHaD, (R.S.)]; Fonds Agir Pour les Maladies Chroniques 2011 (APMC, CDMR R13076CC); AGIRàdom. E.F.S. was supported by the Intramural Research Program of the Eunice Kennedy Shriver National Institute of Child Health and Human Development, National Institutes of Health. C.V. received support from a doctoral grant from University Grenoble Alpes.

The findings expressed in this article are the opinions of the authors and do not necessarily reflect the official position of the Centers for Disease Control and Prevention (CDC). Use of trade names is for identification only and does not imply endorsement by the CDC, the Public Health Service, or the U.S. Department of Health and Human Services.

## References

Barr DB, Wilder LC, Caudill SP, Gonzalez AJ, Needham LL, Pirkle JL. 2005. Urinary creatinine concentrations in the U.S. population: implications for urinary biologic monitoring measurements. Environ Health Perspect 113(2):192–200, PMID: 15687057, https://doi.org/10.1289/ehp.7337.

Bertelsen RJ, Engel SM, Jusko TA, Calafat AM, Hoppin JA, London SJ, et al. 2014. Reliability of triclosan measures in repeated urine samples from Norwegian pregnant women. J Expo Sci Environ Epidemiol 24(5):517–521, PMID: 24472755, https://doi.org/10.1038/jes.2013.95.

Boeniger MF, Lowry LK, Rosenberg J. 1993. Interpretation of urine results used to assess chemical exposure with emphasis on creatinine adjustments: a review. Am Ind Hyg Assoc J 54(10):615–627, PMID: 8237794, https://doi.org/10.1080/15298669391355134.

Braun JM. 2017. Early-life exposure to EDCs: role in childhood obesity and neurodevelopment. Nat Rev Endocrinol 13(3):161–173, PMID: 27857130, https://doi.org/10.1038/nrendo.2016.186.

Braun JM, Kalkbrenner AE, Calafat AM, Bernert JT, Ye X, Silva MJ, et al. 2011. Variability and predictors of urinary bisphenol A concentrations during pregnancy. Environ Health Perspect 119(1):131–137, PMID: 21205581, https://doi.org/10.1289/ehp.1002366.

Braun JM, Smith KW, Williams PL, Calafat AM, Berry K, Ehrlich S, et al. 2012. Variability of urinary phthalate metabolite and bisphenol A concentrations before and during pregnancy. Environ Health Perspect 120(5):739–745, PMID: 22262702, https://doi.org/10.1289/ehp.1104139.

Carroll RJ, Ruppert D, Stefanski, L.A., Craniceanu CM. 2006. *Measurement Error in Nonlinear Models: A Modern Perspective*, Second Edition. Boca Raton, FL: Chapman & Hall/CRC.

CDC (Centers for Disease Control and Prevention). 2015. "Fourth National Report on Human Exposure to Environmental Chemicals, Updated Tables (February, 2015)". Atlanta, GA:U.S. Department of Health and Human Services, Centers for Disease Control and Prevention. https://www.cdc.gov/exposurereport/ [accessed 24 January 2017].

Chen D, Kannan K, Tan H, Zheng Z, Feng YL, Wu Y, et al. 2016. Bisphenol analogues other than BPA: environmental occurrence, human exposure, and toxicity—a review. Environ Sci Technol 50(11):5438–5453, PMID: 27143250, https://doi.org/10.1021/acs.est.5b05387.

Cheung KL, Lafayette RA. 2013. Renal physiology of pregnancy. Adv Chronic Kidney Dis 20(3):209–214, PMID: 23928384, https://doi.org/10.1053/j.ackd.2013.01.012.

Cosmetic Ingredient Review Expert Panel. 2008. Final amended report on the safety assessment of methylparaben, ethylparaben, propylparaben, isopropylparaben, butylparaben, isobutylparaben, and benzylparaben as used in cosmetic products. Int J Toxicol 27 (suppl 4):1–82, PMID: 19101832, https://doi.org/10.1080/10915810802548359.

Crinnion WJ. 2010. The CDC Fourth National Report on Human Exposure to Environmental Chemicals: what it tells us about our toxic burden and how it assist environmental medicine physicians. Altern Med Rev 15(2):101–109, PMID: 20806995.

Dempsey D, Jacob P III, Benowitz NL. 2002. Accelerated metabolism of nicotine and cotinine in pregnant smokers. J Pharmacol Exp Ther 301(2):594–598, PMID: 11961061.

Fisher M, Arbuckle TE, Mallick R, LeBlanc A, Hauser R, Feeley M, et al. 2015. Bisphenol A and phthalate metabolite urinary concentrations: daily and across pregnancy variability. J Expo Sci Environ Epidemiol 25(3):231–239, PMID: 25248937, https://doi.org/10.1038/jes.2014.65.

Guidry VT, Longnecker MP, Aase H, Eggesbø M, Zeiner P, Reichborn-Kjennerud T, et al. 2015. Measurement of total and free urinary phenol and paraben concentrations over the course of pregnancy: assessing reliability and contamination of specimens in the Norwegian Mother and Child Cohort study. Environ Health Perspect 123(7):705–711, PMID: 25782115, https://doi.org/10.1289/ehp.1408325.

HSDB (Hazardous Substances Data Bank). 2016. HSDB. https://toxnet.nlm.nih.gov/cgi-bin/sis/search2 [accessed 24 January 2017].

INERIS (Institut National de l'Environnement Industriel et des Risques). 2014. "Données technico-économiques sur les substances chimiques en France: Bisphénols F et S" [in French]. DRC-14-136881-02238A. http://www.ineris.fr/substances/fr/ [accessed 23 March 2017].

Janjua NR, Mortensen GK, Andersson AM, Kongshoj B, Skakkebaek NE, Wulf HC. 2007. Systemic uptake of diethyl phthalate, dibutyl phthalate, and butyl paraben following whole-body topical application and reproductive and thyroid hormone levels in humans. Environ Sci Technol 41(15):5564–5570, PMID: 17822133.

Jusko TA, Shaw PA, Snijder CA, Pierik FH, Koch HM, Hauser R, et al. 2014. Reproducibility of urinary bisphenol A concentrations measured during pregnancy in the Generation R Study. J Expo Sci Environ Epidemiol 24(5):532–536, PMID: 24736100, https://doi.org/10.1038/jes.2014.23.

Koch HM, Aylward LL, Hays SM, Smolders R, Moos RK, Cocker J, et al. 2014. Inter- and intra-individual variation in urinary biomarker concentrations over a 6-day sampling period. Part 2: personal care product ingredients. Toxicol Lett 231(2):261–269, PMID: 24956590, https://doi.org/10.1016/j.toxlet.2014.06.023.

Krause M, Klit A, Blomberg Jensen M, Søeborg T, Frederiksen H, Schlumpf M, et al. 2012. Sunscreens: are they beneficial for health? An overview of endocrine disrupting properties of UV-filters. Int J Androl 35(3):424–436, PMID: 22612478, https://doi.org/10.1111/j.1365-2605.2012.01280.x.

Lindon JC, Nicholson, JK, Holmes E, eds. 2007. *The Handbook of Metabonomics and Metabolomics*. Amsterdam, Netherlands:Elsevier.

Meeker JD, Cantonwine DE, Rivera-González LO, Ferguson KK, Mukherjee B, Calafat AM, et al. 2013. Distribution, variability, and predictors of urinary concentrations of phenols and parabens among pregnant women in Puerto Rico. Environ Sci Technol 47(7):3439–3447, PMID: 23469879, https://doi.org/10.1021/es400510g.

NLM (National Library of Medicine). 2016. Household Products Database. https://householdproducts.nlm.nih.gov/index.htm [accessed 24 January 2017].

NTP (National Toxicology Program). 2008. NTP Brief on Bisphenol A [CAS NO. 80–05–07]. https://ntp.niehs.nih.gov/ntp/ohat/bisphenol/bisphenol.pdf [accessed 24 January 2017].

Ouidir M, Giorgis-Allemand L, Lyon-Caen S, Morelli X, Cracowski C, Pontet S, et al. 2015. Estimation of exposure to atmospheric pollutants during pregnancy integrating space-time activity and indoor air levels: does it make a difference? Environ Int 84:161–173, PMID: 26300245, https://doi.org/10.1016/j.envint.2015.07.021.

Perrier F, Giorgis-Allemand L, Slama R, Philippat C. 2016. Within-subject pooling of biological samples to reduce exposure misclassification in biomarker-based studies. Epidemiology 27(3):378–388, PMID: 27035688, https://doi.org/10.1097/EDE.0000000000000460.

Persky AM, Eddington ND, Derendorf H. 2003. A review of the effects of chronic exercise and physical fitness level on resting pharmacokinetics. Int J Clin Pharmacol Ther 41(11):504–516, PMID: 14620948.

Philippat C, Wolff MS, Calafat AM, Ye X, Bausell R, Meadows M, et al. 2013. Prenatal exposure to environmental phenols: concentrations in amniotic fluid and variability in urinary concentrations during pregnancy. Environ Health Perspect 121(10):1225–1231, PMID: 23942273, https://doi.org/10.1289/ehp.1206335.

Sandborgh-Englund G, Adolfsson-Erici M, Odham G, Ekstrand J. 2006. Pharmacokinetics of triclosan following oral ingestion in humans. J Toxicol Environ Health Part A 69(20):1861–1873, PMID: 16952905, https://doi.org/10.1080/15287390600631706.

Smith KW, Braun JM, Williams PL, Ehrlich S, Correia KF, Calafat AM, et al. 2012. Predictors and variability of urinary paraben concentrations in men and women, including before and during pregnancy. Environ Health Perspect 120(11):1538–1543, PMID: 22721761, https://doi.org/10.1289/ehp.1104614.

Stacy SL, Eliot M, Etzel T, Papandonatos G, Calafat AM, Chen A, et al. 2017. Patterns, variability, and predictors of urinary triclosan concentrations during pregnancy and childhood. Environ Sci Technol 51(11):6404–6413, PMID: 28516781, https://doi.org/10.1021/acs.est.7b00325.

Völkel W, Colnot T, Csanády GA, Filser JG, Dekant W. 2002. Metabolism and kinetics of bisphenol A in humans at low doses following oral administration. Chem Res Toxicol 15(10):1281–1287, PMID: 12387626.

Wang CS, Yandell BS, Rutledge JJ. 1992. The dilemma of negative analysis of variance estimators of intraclass correlation. Theor Appl Genet 85(1):79–88, PMID: 24197232, https://doi.org/10.1007/BF00223848.

Ye X, Wong LY, Bishop AM, Calafat AM. 2011. Variability of urinary concentrations of bisphenol A in spot samples, first morning voids, and 24-hour collections. Environ Health Perspect 119(7):983–988, PMID: 21406337, https://doi.org/10.1289/ehp.1002701.

Zhou X, Kramer JP, Calafat AM, Ye X. 2014. Automated on-line column-switching high performance liquid chromatography isotope dilution tandem mass spectrometry method for the quantification of bisphenol A, bisphenol F, bisphenol S, and 11 other phenols in urine. J Chromatogr B Analyt Technol Biomed Life Sci 944:152–156, PMID: 24316527, https://doi.org/10.1016/j.jchromb.2013.11.009.