**SUPPLEMENTARY MATERIAL**

**An empirical validation of the within-subject biospecimens pooling approach to minimize exposure misclassification in biomarker-based studies**

**Table of contents**

[**eAppendix 1**. 3](#_Toc521922015)

[**eTable 1.** Characteristics of the 16 pregnant women from the SEPAGES feasibility study included in the current study. 7](#_Toc532441002)

[**eTable 2.** Effect estimates and statistical power to detect associations between biomarker-based exposure to methylparaben and a continuous outcome, as a function of the number of biospecimens collected per subject to assess exposure (1,000 simulation runs with 3,000 subjects each; true effect, $βtrue$ = -100g change in the outcome for each unit increase in the true (unmeasured) exposure). 16](#_Toc532441003)

[**eTable 3.** Effect estimates and statistical power to detect associations between biomarker-based exposure to bisphenol A and a continuous outcome, as a function of the number of biospecimens collected per subject to assess exposure (1,000 simulation runs with 3,000 subjects each; true effect, $βtrue$ = -100g change in the outcome for each unit increase in the true (unmeasured) exposure). 17](#_Toc532441004)

[**eFigure 1.** Daily exposure window – Scatter plots of exposure estimates from Protocol 2 (equal volumes of three spot urines voids were pooled within-subject) against those from Protocol 1 (equal volumes of all spot urine voids collected during the 3 follow-up weeks were pooled within-subject) (mean-centered ln-transformed biomarker concentrations, n = 8 women, N = 20 samples). 8](#_Toc531343072)

[**eFigure 2.** Daily exposure window – Bland-Altman plots for Protocols 1 (equal volumes of all spot urine voids were within-subject pooled) and 2 (equal volumes of three spot urine voids were within-subject pooled) daily averages of ln-transformed biomarkers. 9](#_Toc531343073)

[**eFigure 3.** Weekly exposure window – Scatter plots of exposure estimates from Protocol 2 (equal volumes of three spot urine voids were pooled within-subject for daily pools) against those from Protocol 1 (equal volumes of all spot urine voids were pooled within-subject) (mean-centered ln-transformed biomarker concentrations, n = 8 women, N = 24 samples). 10](#_Toc531343074)

[**eFigure 4.** Weekly exposure window – Bland-Altman plots for Protocols 1 (equal volumes of all spot urine voids were within-subject pooled) and 2 (equal volumes of three spot urine voids were within-subject pooled) weekly averages of ln-transformed biomarkers. 11](#_Toc531343075)

[**eFigure 5.** Pregnancy exposure window – Scatter plots of exposure estimates from Protocols 2-5 against Protocol 1 (pooling of all spot urine samples/day). 12](#_Toc531343076)

[**eFigure 6.** Pregnancy exposure window – Bland-Altman plots for Protocols 3 (A), 4 (B), 5 (C) and 2 (D) against Protocol 1 (pooling of all spot urine samples/day) pregnancy averages of ln-transformed biomarkers. 14](#_Toc531343077)

[**eFigure 7.** Distribution of the simulated exposures in one of our simulated studies for methylparaben (A, ICC = 0.85) and for bisphenol A (B, ICC = 0.38). 18](#_Toc531343078)

[**eFigure 8.** Examples of urinary concentrations of methylparaben (A) and bisphenol A (B) measured with error, for three subjects in one of our simulation runs. 19](#_Toc531343079)

**eAppendix 1.**

**Aim 2: Impact of the within-subject biomarker variability on dose-response functions**

We chose two phenols, methylparaben and bisphenol A, because of their contrasted ICCs, with respectively low and high within-subject variability in the studied population of eight women (pregnancy-specific ICC was 0.85 for methylparaben and 0.38 for bisphenol A).1

*Simulation of exposures*

We simulated a population of 3,000 subjects with one to 20 spot urine samples using a bootstrap method based on the eight women from SEPAGES-feasibility study with the lowest rate of missed urine voids (groups A1 and A2). Each of these women had eight random spot samples assayed for phenol biomarkers among all spot samples collected at three occasions of pregnancy (see Figure 1).

We assigned for each subject $i \left(i=1, …, 3,000\right)$ one pregnant woman out of the eight from the SEPAGES-feasibility study, with, for each biomarker $j (j=1, 2)$ the average (namely $C\_{ij\_{mean}}$) of biomarker ln-concentrations in the eight random spot samples. To obtain unclustered subjects (clusters corresponding to the eight pregnant women), for each participant $i (i=1, …, 3,000)$, and each biomarker 𝑗$ (j=1, 2)$, we generated $X\_{ij}$ the true but unobserved (ln-transformed) exposure from a normal distribution, using a mean (standard deviation, SD) concentration of 4.64 (1.59) µg/L for ln(methylparaben), and 0.91 (0.75) µg/L for ln(bisphenol A). Concentrations values (means and SD) were extracted from the EDEN French mother-child cohort with approximately log-normal distributions,2 and not from the SEPAGES cohort, as phenol biomarker were not yet assessed in the SEPAGES cohort biospecimens when this study was conducted.

Then, for each subject $i$, we generated $k \left(k=1, …, 20\right)$ biospecimens with a bootstrap method, by randomly assigning one out of the eight real random spot samples available (for each subject) to each biospecimen $k$.

For each biospecimen $k$, the biomarker-specific ln-concentration $C\_{ijk}$ was centered around $X\_{ij}$ using the formula $W\_{ijk}=C\_{ijk}+X\_{ij}-C\_{ij\_{mean}}$, where $C\_{ij\_{mean}}$ was the average of biomarker ln-concentrations in the eight random spot samples, $ W\_{ijk}$ corresponded to the resulting $X\_{ij}$-centered biospecimen- and biomarker-specific concentration. We presented examples of $X\_{ij}$, and $W\_{ijk}$ distributions in eFigures 5 and 6.

*Simulation of health outcomes*

For each subject $i$ and each biomarker $j$, a continuous outcome $Y\_{ij}$ was simulated as $Y\_{ij}=β\_{1}X\_{ij}+α+ε\_{ij}$, with $β\_{1}$ the true effect assumed to be -100g by one-unit increase in $X\_{ij}$ and $ε\_{ij}$ the independent normally distributed random error with mean zero. The values for the parameters $α$ (14,900g) and $ε\_{ij}$ (SD, 1,650g) were selected so as to match the distribution of the offspring weight at age 3 years in the French EDEN mother-child cohort 3, to reproduce what was previously done by Perrier *et al.* 4 and because SEPAGES-feasibility stopped at birth and therefore data on offspring weight at 3 years of age were not collected.

We additionally simulated $Y\_{ij}$ considering a null effect of the biomarker (i.e, $β\_{1}=0$) to explore how the risk of type I error was affected.

*Bias and power characterization*

For each chemical $j$, we fitted a linear regression model in the population of 3,000 subjects between the simulated continuous health outcome and a biomarker measurement in one random spot sample ($W\_{ijk}$) and within-subject pools of an increasing number of biospecimens, represented by the average ($\overline{W\_{ij}}$) of ln-transformed biomarker concentrations from two to 20 randomly collected biospecimens, as in the previous theoretical study from Perrier et al.4.

Bias was estimated in percent as the difference between the mean effect estimate ($β$) over 1,000 studies for the surrogates of exposure ($W\_{ijk}$ or $\overline{W\_{ij}}$) and the mean effect estimate for the true unobserved exposure ($β\_{true}$) divided by $β\_{true}$. Negative values of bias correspond to a situation where $β$is lower in absolute value than the true effect estimate $β\_{true}$ (i.e. attenuation) and positive values to a situation where $β$is greater in absolute value than the true effect estimate $β\_{true}$.

Power was calculated as the fraction of the 1,000 studies with a p-value for the association below 0.05.

*A posteriori disattenuation*

We additionally reported *a posteriori* disattenuated effect estimates, obtained by dividing the estimated regression coefficients by the compound-specific ICC.10,11 We used two possible values of the pregnancy-specific ICC: ICC1, corresponding to the value estimated in our study population of eight women (ICC1 was 0.85 and 0.38 for methylparaben and bisphenol A, respectively);1 and ICC2, corresponding to the averaged ICC from previously published studies in pregnant women (0.45 and 0.20 for methylparaben and bisphenol A, respectively).5–10 ICC1 was assumed to correspond to the ideal value, but we also used ICC2 because without repeated assays one cannot estimate ICCs internal to the study population.

**APPENDIX REFERENCES**

1. Vernet C, Philippat C, Calafat AM, et al. Within-Day, Between-Day, and Between-Week Variability of Urinary Concentrations of Phenol Biomarkers in Pregnant Women. *Environ Health Perspect*. 2018;126(3):037005. doi:10.1289/EHP1994

2. Vernet C, Pin I, Giorgis-Allemand L, et al. In Utero Exposure to Select Phenols and Phthalates and Respiratory Health in Five-Year-Old Boys: A Prospective Study. *Environ Health Perspect*. 2017;125(9). doi:10.1289/EHP1015

3. Philippat C, Botton J, Calafat AM, Ye X, Charles M-A, Slama R. Prenatal Exposure to Phenols and Growth in Boys: *Epidemiology*. 2014;25(5):625-635. doi:10.1097/EDE.0000000000000132

4. Perrier F, Giorgis-Allemand L, Slama R, Philippat C. Within-subject pooling of biological samples to reduce exposure misclassification in biomarker-based studies: *Epidemiology*. February 2016:1. doi:10.1097/EDE.0000000000000460

5. Braun JM, Kalkbrenner AE, Calafat AM, et al. Variability and Predictors of Urinary Bisphenol A Concentrations during Pregnancy. *Environ Health Perspect*. 2011;119(1):131. doi:10.1289/ehp.1002366

6. Braun JM, Smith KW, Williams PL, et al. Variability of Urinary Phthalate Metabolite and Bisphenol A Concentrations before and during Pregnancy. *Environ Health Perspect*. 2012;120(5):739-745. doi:10.1289/ehp.1104139

7. Jusko TA, Shaw PA, Snijder CA, et al. Reproducibility of Urinary Bisphenol A Concentrations Measured During Pregnancy in the Generation R Study. *J Expo Sci Environ Epidemiol*. 2014;24(5):532. doi:10.1038/jes.2014.23

8. Meeker JD, Cantonwine DE, Rivera-González LO, et al. Distribution, variability and predictors of urinary concentrations of phenols and parabens among pregnant women in Puerto Rico. *Environ Sci Technol*. 2013;47(7):3439. doi:10.1021/es400510g

9. Philippat C, Wolff MS, Calafat AM, et al. Prenatal Exposure to Environmental Phenols: Concentrations in Amniotic Fluid and Variability in Urinary Concentrations during Pregnancy. *Environ Health Perspect*. 2013;121(10):1225-1231. doi:10.1289/ehp.1206335

10. Smith KW, Braun JM, Williams PL, et al. Predictors and Variability of Urinary Paraben Concentrations in Men and Women, Including before and during Pregnancy. *Environ Health Perspect*. 2012;120(11):1538. doi:10.1289/ehp.1104614

**eTable 1.** Characteristics of the 16 pregnant women from the SEPAGES feasibility study included in the current study.

|  |  |  |  |
| --- | --- | --- | --- |
|  | Groups A1 and A2 (n = 8) |  | Group B (n = 8) |
| Characteristic | No. (%)  |  | No. (%)  |
| **Civil status** Married Cohabitating | 5 (62.5)3 (37.5) |  | 3 (37.5)5 (62.5) |
| **Maternal education** High school or less Up to 3 years of college > 3 years of college Missing | 04 (50)4 (50) |  | 0 6 (75)1 (12.5)1 (12.5) |
| **Smoking history during pregnancy** Yes No | 1 (12.5)7 (87.5) |  | 3 (37.5)5 (62.5) |
| **Parity** 0 1 ≥ 2 | 5 (62.5)2 (25)1 (12.5) |  | 4 (50)3 (37.5)1 (12.5) |
|  | Median (25th, 75th) |  | Median (25th, 75th) |
| **Maternal age at enrolment (years)** | 28.5 (27.0, 31.5) |  | 30.0 (27.5, 33.0) |
| **Gestational age (weeks)** Week 1 of urine collection Week 2 of urine collection Week 3 of urine collection | 14.9 (13.6, 16.2)23.9 (22.9, 25.2)32.4 (31.7, 32.9) |  | 13.1 (11.6, 14.9)22.5 (21.6, 23.4)31.9 (31.7, 32.6) |
| **Time between two successive weeks of urine collection (weeks)** Week 1 – Week 2 Week 2 – Week 3 | 8.9 (8.1, 10.0)8.4 (7.1, 9.3) |  | 9.9 (8.6, 10.8)9.4 (7.4, 11.1) |
| **Number of collected urine voids per day** | 8 (7, 10) |  | 6 (5, 8) |

**eFigure 1.** Daily exposure window – Scatter plots of exposure estimates from Protocol 2 (equal volumes of three spot urines voids were pooled within-subject) against those from Protocol 1 (equal volumes of all spot urine voids collected during the 3 follow-up weeks were pooled within-subject) (mean-centered ln-transformed biomarker concentrations, n = 8 women, N = 20 samples).



The filled red line represents the regression line and the filled black line the identity function. Horizontal and vertical dotted lines indicate the biomarker-specific limit of detection.

**eFigure 2.** Daily exposure window – Bland-Altman plots for Protocols 1 (equal volumes of all spot urine voids were within-subject pooled) and 2 (equal volumes of three spot urine voids were within-subject pooled) daily averages of ln-transformed biomarkers.



The filled red lines represent the biomarker-specific mean differences between Protocols 1 and 2 concentrations, and can be compared to the black dotted line (y=0). Dotted red lines indicate the biomarker-specific upper (mean difference + 1.96 standard deviation [SD]), and lower limits of agreement (mean difference - 1.96 SD).

**eFigure 3.** Weekly exposure window – Scatter plots of exposure estimates from Protocol 2 (equal volumes of three spot urine voids were pooled within-subject for daily pools) against those from Protocol 1 (equal volumes of all spot urine voids were pooled within-subject) (mean-centered ln-transformed biomarker concentrations, n = 8 women, N = 24 samples).



The filled red line represents the regression line and the filled black line the identity function. Horizontal and vertical dotted lines indicate the biomarker-specific limit of detection.

**eFigure 4.** Weekly exposure window – Bland-Altman plots for Protocols 1 (equal volumes of all spot urine voids were within-subject pooled) and 2 (equal volumes of three spot urine voids were within-subject pooled) weekly averages of ln-transformed biomarkers.



The filled red lines represent the biomarker-specific mean differences between Protocols 1 and 2 concentrations, and can be compared to the black dotted line (y=0). Dotted red lines indicate the biomarker-specific upper (mean difference + 1.96 SD), and lower limits of agreement (mean difference - 1.96 SD).

**eFigure 5.** Pregnancy exposure window – Scatter plots of exposure estimates from Protocols 2-5 against Protocol 1 (pooling of all spot urine samples/day).



Protocol 2 corresponds to within-subject pooling of 3 spot urine samples/day over three weeks, Protocol 3 to pregnancy exposure relying on one random spot sample, Protocol 4 on the average of three random spot samples, and Protocol 5 on the average of eight random spot samples. Each point corresponds to pregnancy exposure estimate after mean-centering (ln-transformed concentrations, n = 16 women, N = 16 samples for comparison between Protocols 1 and 2; n = 8 women, N = 8 samples for comparison between Protocol 1 and Protocols 3-5). The filled red line represents the regression line and the filled black line the identity line. Horizontal and vertical dotted lines showed the biomarker-specific limit of detection.

**eFigure 5.** Continued



**eFigure 6.** Pregnancy exposure window – Bland-Altman plots for Protocols 3 (A), 4 (B), 5 (C) and 2 (D) against Protocol 1 (pooling of all spot urine samples/day) pregnancy averages of ln-transformed biomarkers.



Protocol 2 corresponds to within-subject pooling of 3 spot urine samples/day over three weeks, Protocol 3 to pregnancy exposure relying on one random spot sample, Protocol 4 on the average of three random spot samples, and Protocol 5 on the average of eight random spot samples.

The filled red lines represent the biomarker-specific mean differences between Protocols 1 and 2 concentrations, and can be compared to the black dotted line (y=0). Dotted red lines indicate the biomarker-specific upper (mean difference + 1.96 SD), and lower limits of agreement (mean difference - 1.96 SD).

**eFigure 6.** Continued



**eTable 2.** Effect estimates and statistical power to detect associations between biomarker-based exposure to methylparaben and a continuous outcome, as a function of the number of biospecimens collected per subject to assess exposure (1,000 simulation runs with 3,000 subjects each; true effect, $β\_{true}$ = -100g change in the outcome for each unit increase in the true (unmeasured) exposure).

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Number of****biospecimens per subject** |  |  |  | **Within-subject pooling** |  | **Within-subject pooling + *a posteriori* disattenuation using ICC1 *f*** |  | **Within-subject pooling** **+ *a posteriori* disattenuation using ICC2 *g*** |
|  | **True effect** |  | **Estimated effect *a*****(95% CI *b*)** | **Power*c*** | **Bias (%)*d*** |  | **Estimated effect *a* (95% CI *b*)** | **Power*c*** | **Bias (%) *d*** |  | **Estimated effect *a* (95% CI *b*)** | **Power*c*** | **Bias (%) *d*** |
| 1 |  | -100 |  | -71 (-101, -40)*e* | 0.99 | -29 |  | -84 (-119, -47)*e* | 0.99 | -16 |  | -158 (-225, -89)*e* | 0.99 | 58 |
| 2 |  | -100 |  | -83 (-116, -49) | 1 | -17 |  | -91 (-127, -53) | 1 | -9 |  | -134 (-188, -79) | 1 | 34 |
| 3 |  | -101 |  | -89 (-122, -53) | 1 | -12 |  | -94 (-129, -56) | 1 | -7 |  | -125 (-172, -75) | 1 | 24 |
| 4 |  | -99 |  | -91 (-127, -54) | 1 | -9 |  | -95 (-132, -56) | 1 | -5 |  | -118 (-166, -70) | 1 | 19 |
| 5 |  | -100 |  | -93 (-126, -58) | 1 | -7 |  | -96 (-131, -60) | 1 | -4 |  | -116 (-157, -72) | 1 | 15 |
| 6 |  | -101 |  | -94 (-129, -60) | 1 | -6 |  | -97 (-133, -62) | 1 | -3 |  | -113 (-155, -72) | 1 | 13 |
| 7 |  | -99 |  | -94 (-129, -57) | 1 | -5 |  | -96 (-132, -58) | 1 | -3 |  | -110 (-151, -67) | 1 | 11 |
| 8 |  | -99 |  | -95 (-130, -59) | 1 | -5 |  | -97 (-133, -61) | 1 | -3 |  | -109 (-150, -68) | 1 | 10 |
| 9 |  | -102 |  | -96 (-134, -58) | 1 | -4 |  | -98 (-136, -59) | 1 | -2 |  | -109 (-152, -65) | 1 | 9 |
| 10 |  | -100 |  | -96 (-133, -60) | 1 | -4 |  | -98 (-135, -61) | 1 | -2 |  | -108 (-149, -67) | 1 | 8 |
| 12 |  | -101 |  | -97 (-134, -61) | 1 | -3 |  | -99 (-136, -62) | 1 | -2 |  | -107 (-148, -68) | 1 | 7 |
| 15 |  | -100 |  | -97 (-135, -61) | 1 | -3 |  | -98 (-136, -61) | 1 | -2 |  | -105 (-145, -65) | 1 | 5 |
| 18 |  | -100 |  | -98 (-135, -61) | 1 | -2 |  | -99 (-137, -61) | 1 | -1 |  | -105 (-144, -65) | 1 | 4 |
| 20 |  | -99 |  | -97 (-132, -60) | 1 | -2 |  | -98 (-133, -61) | 1 | -1 |  | -103 (-140, -64) | 1 | 4 |
| *a* Mean of effect estimates over 1,000 simulated studies.*b* Empirical confidence interval, corresponding to the empirical 2.5 and 97.5 percentiles of the health effect estimates over the 1,000 simulation runs.*c* Statistical power, estimated as the proportion of studies in which the *P-*value of the parameter characterizing the association between the error-prone exposure variables ($W\_{ij}$)and the continuous outcome was below 0.05.*d* Difference between the true effect and the effect estimate, divided by the true effect.*e* Corresponds to a situation without pooling.*f*ICC1 = 0.85 and corresponds to the value estimated in our study population of eight women. *g* ICC2 = 0.45 and corresponds to the average from previously published studies in pregnant women.15–20 |

**eTable 3.** Effect estimates and statistical power to detect associations between biomarker-based exposure to bisphenol A and a continuous outcome, as a function of the number of biospecimens collected per subject to assess exposure (1,000 simulation runs with 3,000 subjects each; true effect, $β\_{true}$ = -100g change in the outcome for each unit increase in the true (unmeasured) exposure).

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  |  |  |  | **Within-subject pooling** |  | **Within-subject pooling + *a posteriori* disattenuation using ICC1*f*** |  | **Within-subject pooling** **+ *a posteriori* disattenuation using ICC2 *g*** |

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Number of biospecimens per subject** |  | **True Effect** |  | **Estimated effect *a* (95% CI *b*)** | **Power*c*** | **Bias (%) *d*** |  | **Estimated effect *a*****(95% CI *b*)** | **Power*c*** | **Bias (%) *d*** |  | **Estimated effect *a*****(95% CI *b*)** | **Power*c*** | **Bias (%) *d*** |
| 1 |  | -100 |  | -31 (-76, 16) *e* | 0.27 | -69 |  | -81 (-199, 43) *e* | 0.27 | -19 |  | -153 (-378, 82) *e* | 0.27 | 54 |
| 2 |  | -100 |  | -48 (-98, 4) | 0.40 | -52 |  | -86 (-178, 7) | 0.40 | -13 |  | -143 (-294, 12) | 0.40 | 43 |
| 3 |  | -101 |  | -59 (-121, -0) | 0.48 | -42 |  | -91 (-187, -0) | 0.48 | -10 |  | -137 (-283, -0) | 0.48 | 36 |
| 4 |  | -99 |  | -65 (-129, 1) | 0.53 | -34 |  | -92 (-182, 2) | 0.53 | -7 |  | -130 (-259, 2) | 0.53 | 31 |
| 5 |  | -100 |  | -70 (-132, -5) | 0.54 | -30 |  | -93 (-175, -6) | 0.54 | -8 |  | -126 (-238, -8) | 0.54 | 25 |
| 6 |  | -101 |  | -74 (-140, -8) | 0.60 | -27 |  | -94 (-178, -10) | 0.60 | -7 |  | -123 (-233, -13) | 0.60 | 22 |
| 7 |  | -99 |  | -75 (-142, 1) | 0.58 | -24 |  | -92 (-175, 2) | 0.58 | -7 |  | -118 (-223, 2) | 0.58 | 19 |
| 8 |  | -99 |  | -78 (-145, -5) | 0.60 | -21 |  | -94 (-175, -7) | 0.60 | -5 |  | -117 (-218, -8) | 0.60 | 18 |
| 9 |  | -102 |  | -81 (-152, -7) | 0.64 | -20 |  | -96 (-179, -8) | 0.64 | -5 |  | -118 (-219, -10) | 0.64 | 16 |
| 10 |  | -100 |  | -82 (-152, -5) | 0.62 | -18 |  | -96 (-177, -6) | 0.62 | -4 |  | -115 (-213, -7) | 0.62 | 15 |
| 12 |  | -101 |  | -86 (-157, -16) | 0.65 | -15 |  | -97 (-179, -18) | 0.65 | -4 |  | -114 (-210, -22) | 0.65 | 13 |
| 15 |  | -100 |  | -87 (-159, -13) | 0.64 | -13 |  | -97 (-177, -15) | 0.64 | -3 |  | -110 (-202, -17) | 0.64 | 11 |
| 18 |  | -100 |  | -90 (-166, -16) | 0.66 | -10 |  | -98 (-181, -18) | 0.66 | -2 |  | -110 (-202, -20) | 0.66 | 9 |
| 20 |  | -99 |  | -89 (-162, -14) | 0.63 | -10 |  | -96 (-175, -15) | 0.63 | -3 |  | -106 (-194, -17) | 0.63 | 7 |
| *a* Mean of effect estimates over 1,000 simulated studies.*b* Empirical confidence interval, corresponding to the empirical 2.5 and 97.5 percentiles of the health effect estimates over the 1,000 simulation runs.*c* Statistical power, estimated as the proportion of studies in which the *P-*value of the parameter characterizing the association between the error-prone exposure variables ($W\_{ij}$)and the continuous outcome was below 0.05.*d* Difference between the true effect and the effect estimate, divided by the true effect.*e* Corresponds to a situation without pooling.*f* ICC1 = 0.38 and corresponds to the value estimated in our study population of eight women. *g* ICC2 = 0.20 and corresponds to the average from previously published studies in pregnant women.15–20 |

**eFigure 7.** Distribution of the simulated exposures in one of our simulated studies for methylparaben (A, ICC = 0.85) and for bisphenol A (B, ICC = 0.38).

1. Methylparaben



1. Bisphenol A



$X$ is the *true* unobserved average exposure and $W$ is the surrogate exposure measured with error (using one spot sample).

**eFigure 8.** Examples of urinary concentrations of methylparaben (A) and bisphenol A (B) measured with error, for three subjects in one of our simulation runs.

|  |  |
| --- | --- |
| 1. Methylparaben (ICC = 0.85)
 | 1. Bisphenol A (ICC = 0.38)
 |



The dashed lines display the true unobserved exposure average of urinary ln-concentrations of the corresponding chemical over a toxicologically relevant exposure window.