

Supplementary Material S5: Exploring determinants of BPA exposure from combined population.

// Approach for exploring the determinants of urine BPA levels //

Statistical analysis

In the regression model, the dependent variable BPA was log-transformed for the investigation of factors associated with BPA exposure levels to ensure that the distribution of the residuals was closer to a normal distribution. Collinearity between the selected variables was checked to avoid including variables that are highly correlated with each other in the regression model. As a result, the variable 'laboratory analytical method' was retained instead of the variable 'type of sampling/storage tube'. The variable 'sampling year' was available in all studies and was preferred to the variable 'samplingmonth' (due to competition in the models, although these two variables are not identical in a vacuum). As the dataset included observations from different European studies, analyses were performed using Linear Mixed-Effects Regression (LMER) models, which can include random effects. Random effects were used to model the variability between studies. The LMER model also took into account the hierarchical structure of the grouped individuals and the variability within and between studies.

In order to perform analyses on the most complete data set possible, it was decided to impute the explanatory variables. Variables were imputed if they met the following two conditions: at least three studies had collected the variable, and the variable had less than 50% missing data in each of the contributing studies.

Imputation was performed using R, the VIM package and the "matchImpute" command. The "matchImpute" command searches for similar observations in the dataset using a matching method. The matching variables chosen for our imputations were "ageyears", "height" and "education (Isced levels)". These variables were common to all studies and had very few missing data. If one of these variables was missing, an intermediate imputation was performed using the other two matching variables and the variable "urinary creatinine" to ensure that the three initial matching variables were not missing. Once the matched observations were identified, "matchImpute" used least squares regression methods to predict missing values based on the known values of the matched observations. The model included the dependent variable BPA, the independent variable of interest, and the random effect, which in this study was the "study collection".

The model estimation method used was Restricted Maximum Likelihood (REML), which is used to estimate the parameters of linear mixed-effects models, emphasizing fixed effects and minimizing the influence of random effects. For models with multiple independent variables, model selection was performed based on variables that were individually significant, variables with fewer missing values, and collinearity between variables. All models were systematically adjusted for the variables "ageyears", "urinary creatinine" and "sampling year".

Finally, in order to take into account the urine dilution, we fitted two similar models to either the creatinine or to the specific gravity.

All statistical analyses were performed using R 4.1.3 (R Core Team).

Eligible variables for building models

Figure S8 shows all variables for which it was possible to investigate the variation of total BPA levels, based only on studies with sufficient data (see **Table S2**). The variables considered are those for which at least 3 studies had at least 50% of the data available (the 50% missing data should therefore be imputed for the analysis).

Table S3 shows the results of the bivariate analyses performed with each of the eligible variables after the imputation process. The bivariate regressions were only performed on the total population from the studies that had all collected the explanatory variable under analysis. For example, the variable "agegroup" concerned 4058 women (the total number of women, including 14/14 studies that collected the variable), whereas the variable "pregnancy" concerned only 3454 women (the total number of women, including the 9/14 studies that collected the variable) and the variable "gaweeek_m" only concerned 549 women (the total number of women, including the 3/14 studies that collected the variable).

Several linear-mixed regression models were then explored separately with each of the variables that appeared significant in the bivariate analyses with a set of contributing studies. All models were systematically adjusted for a set of confounding variables that were present in all the contributing studies: urine creatinine (linear variable, including as covariate), ageyears (linear variable), sampling year (linear variable), BMI (ordinal variable), and lab analytical method (categorical variable). In all of the models examined with the entire combined population, no variable other than urinary creatinine was significantly associated with BPA levels.

Comparative analysis models including "specific gravity" instead of "urinary creatinine" could only be performed with the three studies in which both these variables were available (3xG_BE, DEMOCOPHES_SE and FLESHIII_BE). In these models, no final variable appeared to be significantly associated with BPA levels, except for the significant association with urinary creatinine in the model including urinary creatinine.

Variable ^a	N obs ^b	N ^c	3XG	DEMOCOPHES	DEMOCOPHES	DEMOCOPHES	DEMOCOPHES	DEMOCOPHES	DEMOCOPHES	ELFE	ESB	FLEHSIII	HELIX-BIB	HELIX-RHEA	IBS	OCC
			(Belgium) n = 150	(Belgium) n = 129	(Denmark) n = 145	(Spain) n = 115	(Luxembourg) n = 58	(Sweden) n = 98	(Slovenia) n = 107	(France) n = 1741	(Germany) n = 60	(Belgium) n = 101	(UK) n = 205	(Greece) n = 194	(Israel) n = 117	(Denmark) n = 838
specific_gravity	358	358														
urine_creatinine	4058	4058														
age_years	4058	4058														
sampling_year	4058	4058														
matrix	4058	4058														
tube	4058	4058														
lab_analytical_method	4058	4058														
population_type	4058	4058														
country	4058	4058														
height ^d	3827	4058														
weight ^e	3653	3908														
sampling_month	3859	3864														
birthplace	1079	1080														
pregnancy	3454	3454														
gaweeks_m	503	549														
iscd	2962	3220														
occupation	552	652														
occupation_status	776	813														
degurba	1763	1801														
nuts1	2704	2704														
nuts2	963	963														
smoking_status	3809	4058														
smoking_cigday ^f	2044	2290														
passive_smoking	2917	3220														
personal_care2	1042	1080														
plastic_floor	1874	2393														
water_tap	2024	2494														
food_packaged	2031	2393														
food_canned	2008	2345														

Fig. S8. Overview of variables eligible for regression analysis to investigate determinants of BPA levels based on data availability from contributing studies. ^a Variables considered are those for which at least 3 studies had at least 50% of the data available (the 50% missing data should therefore be imputed for the analysis). Data imputation per data collection for each variable from the original contributing studies: i) cells in red, variable of interest not imputed because not collected by the original contributing study or variable with more than 50% of missing data from the original contributing study; ii) cells in orange, variable of interest imputed: this concerns only the variables with the half or less than 50% of missing data from the original contributing studies; iii) cells in green, variable of interest with complete data for all participants from the contributing study. ^b Total number of observations from the eligible contributing studies with data available for the variable of interest before imputation. The PBAT study (n= 162) was not been included in the analysis due to several factors, including the high LOQ in the study and the low level of quantification of total BPA. ^c Total number of participants with data after the imputation process. ^d Variable used to calculate BMI (after imputation). ^e Variable used for BMI calculation (after imputation). ^f Variable combining information from current smoking status (yes/no) and current number of cigarettes consumed by smokers.

Table S3. Variation of total urine BPA concentrations in µg/L by analytical parameters and sociodemographic characteristics. Bivariate linear mixed regressions for combined population with contributing studies as random factor per each variable of interest. The change of total urine BPA equals exp (beta).

Variable	Category	N ^a	Beta ^b	Beta (95% CI)		p value	Contributing studies ^c
				LL	UL		
Urine creatinine (linear)		4058	1.453917	1.421360	1.487313	< 0.0001	All, except PBAT
Specific gravity (linear)		358	1.6646926	0.6669279	4.273144	ns	3XG; DEMOCOPHES; and FLEHSIII
Age years (linear)		4058	0.9953052	0.9922958	0.9983846	0.002	All, except PBAT
Age years	< 30	1532	Reference = 0				
	30-39	1911	0.9299584	0.8973751	0.964164	0.02	
	40-49	471	0.9622526	0.8998336	1.031133	ns	All, except PBAT
	50-59	81	0.9224489	0.7756464	1.096414	ns	
	>= 60	63	0.8763336	0.7173775	1.0694	ns	
Sampling year (linear)		4058	0.9365773	0.9041837	0.9699629	0.0002	All, except PBAT
Sampling year	2007	277	Reference = 0				
	2008	152	0.8308683	0.7396843	0.9295414	0.02	
	2009	30	0.8329658	0.6363434	1.0775121	ns	
	2010	7	0.3823688	0.2475415	0.5866505	0.004	All, except PBAT
	2011	2928	0.8919434	0.7222498	1.1016221	ns	
	2012	523	0.7944686	0.6408422	0.9848061	ns	
	2013	40	0.7757133	0.5957502	1.007129	ns	
	2014	101	0.7131951	0.4979769	1.0217103	ns	
Sampling month	January	279	Reference = 0				
	February	102	0.9299346	0.8221664	1.0519233	ns	
	March	107	0.8854573	0.7816356	1.0028099	ns	
	April	117	1.0343495	0.9158021	1.168055	ns	3xG; DEMOCOPHES_BE;
	May	136	0.9659017	0.8598088	1.0847822	ns	DEMOCOPHES_DK; DEMOCOPHES_ES;
	June	718	0.9771576	0.8910657	1.0710468	ns	DEMOCOPHES_LU; DEMOCOPHES_SE;
	July	112	0.9571982	0.8461558	1.0825316	ns	DEMOCOPHES_SI; ELFE; ESB; FLESHIII;
	August	119	0.997125	0.8839056	1.1245398	ns	HELIX_BIB; IBS; and OCC

	September	732	0.9124582	0.8313087	1.0010965	<i>ns</i>	
	October	201	0.8963765	0.8066387	0.9976752	<i>ns</i>	
	November	833	0.938194	0.8587882	1.0250227	<i>ns</i>	
	December	291	1.0050057	0.916139	1.1047853	<i>ns</i>	
Matrix	Urine-spot (US)	2358	1.115535	0.757817	1.642279	<i>ns</i>	
	Urine-morning (UM)	1640	1.147299	0.7880516	1.670462	<i>ns</i>	All, except PBAT
	Urine-24h (UD)	60	Reference = 0				
Sampling/storage tube	Polypropylene	1218	1.103091	0.90373166	1.346124	<i>ns</i>	
	Glass	2840	Reference = 0				All, except PBAT
Lab analytical method	GC-MS	101	0.7278259	0.5030678	1.053045	<i>ns</i>	
	GC-MS/MS	1965	0.8501986	0.6473299	1.114034	<i>ns</i>	
	HPLC-MS/MS	175	0.8902869	0.6584974	1.205495	<i>ns</i>	All, except PBAT
	LC-MS/MS	1289	0.8496968	0.6617688	1.089741	<i>ns</i>	
	Online SPE-HPLC-MS/MS	399	Reference = 0				
Type of population	Pregnant women	3128	0.8943152	0.7452329	1.074481	<i>ns</i>	
	General population	930	Reference = 0				All, except PBAT
Country	Belgium (BE)	380	1.3165038	1.0244231	1.695814	<i>ns</i>	
	Denmark (DK)	983	1.2875205	0.9807195	1.667829	<i>ns</i>	
	Greece (EL)	194	1.6600032	1.2177333	2.262901	<i>ns</i>	
	Germany (DE)	60	1.1893817	0.8585688	1.647659	<i>ns</i>	
	Israel (IL)	117	1.6339653	1.1929101	2.238092	<i>ns</i>	
	Luxembourg (LU)	58	1.414201	1.0200767	1.96071	<i>ns</i>	All, except PBAT
	Slovenia (SI)	107	1.3364956	0.9746424	1.832693	<i>ns</i>	
	Spain (ES)	115	1.521229	1.1103727	2.084109	<i>ns</i>	
	Sweden (SE)	98	1.2329439	0.8980461	1.692731	<i>ns</i>	
	United Kingdom (UK)	205	1.3862281	1.0172978	1.888953	<i>ns</i>	
	France (FR)	1741	Reference = 0				
Region	Northern Europe	1286	1.040014	0.8600506	1.257844	<i>ns</i>	
	Southern Europe	416	1.204984	0.9776076	1.486171	<i>ns</i>	All, except PBAT
	Western Europe	2239	Reference = 0		1.806394		

	Outside Europe	117	1.310748	0.9515352	1.806394	<i>ns</i>	
Body mass index (BMI)	Underweight	65	1.020279	0.8966204	1.161279	<i>ns</i>	DEMOCOPHES_BE; DEMOCOPHES_DK;
	Overweight	1321	1.044154	1.0029421	1.086734	<i>ns</i>	DEMOCOPHES_ES; DEMOCOPHES_LU;
	Obesity	829	1.100326	1.0502607	1.152385	0.03	DEMOCOPHES_SE; DEMOCOPHES_SI;
	Normal	1693	Reference = 0			<i>ns</i>	ELFE; ESB; FLESHIII; HELIX_BIB;
Birthplace	North America	9	0.8216099	0.596645	1.125764	<i>ns</i>	HELIX_RHEA; IBS; and OCC
	South/Central America	9	1.1309472	0.8246591	1.54432	<i>ns</i>	3xG; DEMOCOPHES_BE;
	Africa	74	1.0527907	0.8962929	1.228595	<i>ns</i>	DEMOCOPHES_DK; DEMOCOPHES_ES;
	Asia/Middle East	129	1.0275184	0.8676538	1.219911	<i>ns</i>	DEMOCOPHES_LU; DEMOCOPHES_SE;
	Europe	988	Reference = 0				DEMOCOPHES_SI; ESB; FLESHIII; and IBS
Education (ISCED level)	Low education: isced 0-2	273	Reference = 0				3xG; DEMOCOPHES_BE;
	Medium education: isced 3-4	822	1.0127385	0.9428202	1.088068	<i>ns</i>	DEMOCOPHES_DK; DEMOCOPHES_ES;
	High education: isced ≥ 5	2125	0.9414242	0.8801035	1.006989	<i>ns</i>	DEMOCOPHES_LU; DEMOCOPHES_SE;
Pregnancy	Yes	3130	0.962086	0.7696111	1.191731	<i>ns</i>	DEMOCOPHES_SI; ELFE; ESB; FLESHIII;
	No	324	Reference = 0				HELIX_BIB; HELIX_RHEA; and OCC
Gestational weeks (linear)		549	0.9902441	0.9816949	0.9949556	<i>ns</i>	3XG; HELIX_BIB; and HELIX_RHEA
Occupation status	Full time	392	Reference = 0				
	Part time	219	0.9109950	0.8414654	0.986584	0.03	
	Unemployed	36	0.9279076	0.7927403	1.086681	<i>ns</i>	DEMOCOPHES_BE; DEMOCOPHES_DK;
	Pupil, student	76	0.9333168	0.7674089	1.128664	<i>ns</i>	DEMOCOPHES_ES; DEMOCOPHES_LU;
	Inactive person and other ^c	36	1.0237933	0.8460272	1.220762	<i>ns</i>	DEMOCOPHES_SE; DEMOCOPHES_SI;
	In compulsory military	7	1.1157297	0.7720974	1.583414	<i>ns</i>	ESB; and FLESHIII
	Fulfilling domestic tasks	50	1.0323537	0.8954423	1.194289	<i>ns</i>	
Degree of urbanization	City	402	Reference = 0				3xG; DEMOCOPHES_BE;
	Town or suburb	1212	1.021057	0.9252792	1.122585	<i>ns</i>	DEMOCOPHES_DK; DEMOCOPHES_ES;
	Rural area	187	1.079207	0.9735486	1.197349	<i>ns</i>	DEMOCOPHES_LU; DEMOCOPHES_SE;
NUTS1	BE1	62	1.5200491	1.4619173	1.9256671	0.02	DEMOCOPHES_SI; ESB; FLESHIII; and OCC
	BE2	278	1.1657715	1.0392053	1.2403712	<i>ns</i>	3xG; DEMOCOPHES_BE;
	BE3	40	1.3726054	1.2854087	1.7851906	<i>ns</i>	DEMOCOPHES_DK; DEMOCOPHES_ES;
	DEA	60	1.1167800	0.9711282	1.2842768	<i>ns</i>	DEMOCOPHES_LU; DEMOCOPHES_SE;

	DK0	145	1.4589599	1.3153430	1.6182578	0.04	
	ES3	58	1.3993450	1.2146459	1.6121294	<i>ns</i>	
	ES4	57	1.4585240	1.2648235	1.6818886	0.04	
	FR1	193	Reference = 0				
	FRB	96	0.9360277	0.8319221	1.0531608	<i>ns</i>	
	FRC	170	0.9534535	0.8633240	1.0529924	<i>ns</i>	
	FRD	39	0.8649901	0.7328750	1.0209215	<i>ns</i>	
	FRE	316	0.9190260	0.8430856	1.0018067	<i>ns</i>	
	FRF	204	1.0068720	0.9158063	1.1069931	<i>ns</i>	
	FRG	123	0.9069565	0.8133592	1.0113245	<i>ns</i>	
	FRH	134	0.8542130	0.7681803	0.9498809	0.01	
	FRI	163	0.8978577	0.8120677	0.9927110	0.05	
	FRJ	171	0.9837651	0.8909078	1.0863006	<i>ns</i>	
	FRK	132	0.9216172	0.8284035	1.0253196	<i>ns</i>	
	LU0	58	1.3279127	1.1526419	1.5298350	<i>ns</i>	
	SE1	47	1.1572469	0.9924564	1.3493998	<i>ns</i>	
	SE3	51	1.1580852	0.9980212	1.3438204	<i>ns</i>	
	SI0	107	1.2549138	1.1194828	1.4067288	<i>ns</i>	
Smoking	Smoker	573	1.040982	0.9936646	1.09075	<i>ns</i>	All, except PBAT
	Non-smoker	3485	Reference = 0				
Smoking cig/day	Smoker, 1-9 cig./day	320	1.0368767	0.9762341	1.101376	<i>ns</i>	3xG; ELFE; HELIX_BIB; and HELIX_RHEA
	Smoker, 10-19 cig./day	100	0.9425406	0.8499648	1.045424	<i>ns</i>	
	Smoker, ≥ 20 cig./day	34	1.0584048	0.8865498	1.264002	<i>ns</i>	
	No smoker	3536	Reference = 0				
Passive smoking	Yes	1282	1.057477	1.016459	1.100775	0.01	DEMOCOPHES_BE; DEMOCOPHES_DK; DEMOCOPHES_ES; DEMOCOPHES_LU; DEMOCOPHES_SE; DEMOCOPHES_SI; ELFE; ESB; FLESHIII; HELIX_BIB; HELIX_RHEA; IBS; and OCC
	No	2776	Reference = 0				
Plastic floor	Yes	525	1.039477	0.989667	1.09129	<i>ns</i>	DEMOCOPHES_BE; DEMOCOPHES_DK; DEMOCOPHES_ES; DEMOCOPHES_LU; DEMOCOPHES_SE; DEMOCOPHES_SI; and ELFE
	No	1868	Reference = 0				

Water tap consumption	Yes	2013	0.9840775	0.9348273	1.0365	<i>ns</i>	DEMOCOPHES_BE; DEMOCOPHES_DK; DEMOCOPHES_ES; DEMOCOPHES_LU; DEMOCOPHES_SE; DEMOCOPHES_SI; ELFE; and IBS
	No	481	Reference = 0				
Food packaged (linear)		2393	0.9993341	0.9871379	1.011464	<i>ns</i>	DEMOCOPHES_BE; DEMOCOPHES_DK; DEMOCOPHES_ES; DEMOCOPHES_LU; DEMOCOPHES_SE; DEMOCOPHES_SI; and ELFE
Food packaged	1	1125	1.009926	0.9473782	1.077871	<i>ns</i>	
	2	678	1.017164	0.9529275	1.08573	<i>ns</i>	DEMOCOPHES_BE; DEMOCOPHES_DK; DEMOCOPHES_ES; DEMOCOPHES_LU; DEMOCOPHES_SE; DEMOCOPHES_SI; and ELFE
	3	45	1.145296	0.9752084	1.34807	<i>ns</i>	
	4	214	0.965659	0.8678566	1.077752	<i>ns</i>	
	0	331	Reference = 0				
Food canned (linear)		2345	1.013119	1.000876	1.02548	0.04	DEMOCOPHES_BE; DEMOCOPHES_DK; DEMOCOPHES_ES; DEMOCOPHES_SE; ELFE; and IBS
Food canned	1	1293	0.9694467	0.8953924	1.048174	<i>ns</i>	
	2	547	1.0171873	0.9339295	1.105748	<i>ns</i>	DEMOCOPHES_BE; DEMOCOPHES_DK; DEMOCOPHES_ES; DEMOCOPHES_SE; ELFE; and IBS
	3	112	1.0732253	0.9436001	1.221196	<i>ns</i>	
	4	119	1.0362384	0.9102881	1.179191	<i>ns</i>	
	0	274	Reference = 0				

^a Total number of participants.

^b Beta is the value of the log-transformed change in BPA (coefficient from a linear regression mixed effects model, fitted with a random intercept for the contributing study).

^c Contributing studies = 3xG (Belgium); DEMOCOPHES studies (Belgium, Denmark, Spain, Luxembourg, Sweden, and Slovenia); ELFE (France); ESB (Germany); FLEHS (Belgium); HELIX-BIB (United Kingdom); HELIX-RHEA (Greece); IBS (Israel); OCC (Denmark); and PBAT (Austria).

73

74

75

76