

A Demonstration of the Uncertainty in Predicting the Estrogenic Activity of Individual Chemicals and Mixtures From an *In Vitro* Estrogen Receptor Transcriptional Activation Assay (T47D-KBluc) to the *In Vivo* Uterotrophic Assay Using Oral Exposure

Justin M. Conley,^{*,1} Bethany R. Hannas,^{*,†,1} Johnathan R. Furr,^{*,‡}
Vickie S. Wilson,^{*} and L. Earl Gray Jr.^{*,2}

^{*}Toxicity Assessment Division, U.S. Environmental Protection Agency, Research Triangle Park, North Carolina 27711; [†]Dow Chemical Company, Midland, Michigan 48674; and [‡]Southern Research, Birmingham, Alabama 35205

¹J.M.C. and B.R.H. contributed equally to this work.

²To whom correspondence should be addressed at USEPA, B105-04, 109 TW Alexander Dr., Research Triangle Park, NC 27711. Fax: (919) 541-4017. E-mail: gray.earl@epa.gov.

ABSTRACT

In vitro estrogen receptor assays are valuable tools for identifying environmental samples and chemicals that display estrogenic activity. However, *in vitro* potency cannot necessarily be extrapolated to estimates of *in vivo* potency because *in vitro* assays are currently unable to fully account for absorption, distribution, metabolism, and excretion. To explore this issue, we calculated relative potency factors (RPF), using 17 α -ethinyl estradiol (EE2) as the reference compound, for several chemicals and mixtures in the T47D-KBluc estrogen receptor transactivation assay. *In vitro* RPFs were used to predict rat oral uterotrophic assay responses for these chemicals and mixtures. EE2, 17 β -estradiol (E2), benzyl-butyl phthalate (BBP), bisphenol-A (BPA), bisphenol-AF (BPAF), bisphenol-C (BPC), bisphenol-S (BPS), and methoxychlor (MET) were tested individually, while BPS + MET, BPAF + MET, and BPAF + BPC + BPS + EE2 + MET were tested as equipotent mixtures. *In vivo* ED₅₀ values for BPA, BPAF, and BPC were accurately predicted using *in vitro* data; however, E2 was less potent than predicted, BBP was a false positive, and BPS and MET were 76.6 and 368.3-fold more active *in vivo* than predicted from the *in vitro* potency, respectively. Further, mixture ED₅₀ values were more accurately predicted by the dose addition model using individual chemical *in vivo* uterotrophic data (0.7–1.5-fold difference from observed) than *in vitro* data (1.4–86.8-fold). Overall, these data illustrate the potential for both underestimating and overestimating *in vivo* potency from predictions made with *in vitro* data for compounds that undergo substantial disposition following oral administration. Accounting for aspects of toxicokinetics, notably metabolism, in *in vitro* models will be necessary for accurate *in vitro*-to-*in vivo* extrapolations.

Key words: mixture; T47D-KBluc; bisphenols; estrogen; uterotrophic; oral exposure.

The Endocrine Disruptor Screening Program (EDSP) was established by the U.S. Environmental Protection Agency (USEPA) to screen and test chemicals for endocrine disrupting effects in response to directives of the Food Quality Protection Act (FQPA) and amendments of the Safe Drinking Water Act (SDWA) (Juberg *et al.*, 2014). EDSP takes a tiered approach to screening and further testing of compounds for estrogen, androgen, and/or thyroid hormone activity; however, a backlog of thousands of compounds remains eligible for screening (Judson *et al.*, 2009). The development and use of high throughput screening (HTS) *in vitro* assays is becoming more prevalent as the backlog of compounds requiring toxicity testing, as well as the desire to reduce laboratory animal use, increases. Despite the urgency associated with this endeavor, thorough validation and awareness of both the advantages and current limitations of *in vitro* assays for making predictions of *in vivo* responses are critical to their successful implementation.

Many *in vitro* assays detect (ant)agonistic or binding properties of an endocrine active compound to a particular target receptor. These assays are valuable in determining the mechanism of action and potency that particular compounds might have at the receptor level. Nevertheless, *in vitro* assays are not yet able to fully account for absorption, distribution, metabolism, and excretion (ADME), or other systemic toxicities of a test compound, which may lead to false negative or false positive assay results (Coecke *et al.*, 2006). For that reason, anchoring *in vitro* results to *in vivo* results is an essential exercise in the validation process, to account for biological processes that are not represented by *in vitro* assays.

Numerous synthetic compounds have been categorized as xenoestrogens and are present as contaminants in various environmental media, including sources of drinking water (Kuch and Ballschmiter, 2001). Disruption of estrogen-regulated processes by xenoestrogens can have permanent, deleterious

consequences on growth, differentiation, and reproduction, especially during sensitive developmental stages (Foster and Gray, 2013). In addition, the estrogen receptor (ER) is considered the most “promiscuous” nuclear hormone receptor due to its propensity to bind a broad range of ligands, the majority of which display agonist activity (Blair *et al.*, 2000). Given the structural diversity of ER ligands, the estrogenicity of many compounds has been difficult to accurately predict based on *in silico* structure–activity relationships (SAR) and metabolic activation pathways (Elsby *et al.*, 2000). As such, it is a priority to guarantee the highest confidence in both *in vitro* and *in vivo* screening assay results for these compounds to protect human and wildlife health from exposure to adverse levels of xenoestrogens. Further, the evaluation of chemical estrogenic activity is currently required by the FQPA and SDWA and is therefore one of the most prevalent types of endocrine action screened for by EDSP. For these reasons, the current study focuses on two estrogen-responsive assays, one *in vitro* and one *in vivo*, with the goal of gaining greater understanding of the strengths and limitations of the former.

We selected four xenoestrogens and one endogenous estrogen that display a broad range of potencies due to a diversity of *in vivo* ADME [bisphenol A (BPA), benzyl butyl phthalate (BBP), 17 α -ethinyl estradiol (EE2), 17 β -estradiol (E2), and methoxychlor (MET)] and three xenoestrogens with relatively unknown *in vivo* potency or ADME [bisphenol C (BPC), bisphenol AF (BPAF), and bisphenol S (BPS)] to evaluate *in vitro* versus oral *in vivo* potencies (Figure 1). EE2 was used as the reference chemical for calculating relative potency factors (RPF) because it is one of the most potent, orally effective estrogenic chemicals. Conversely, BBP was included because it is at the opposite end of the spectrum, undergoing rapid and extensive metabolism and excretion. MET is a prototypical “proestrogen”, in that the parent compound is bioactivated to more potent metabolites. In addition, given the

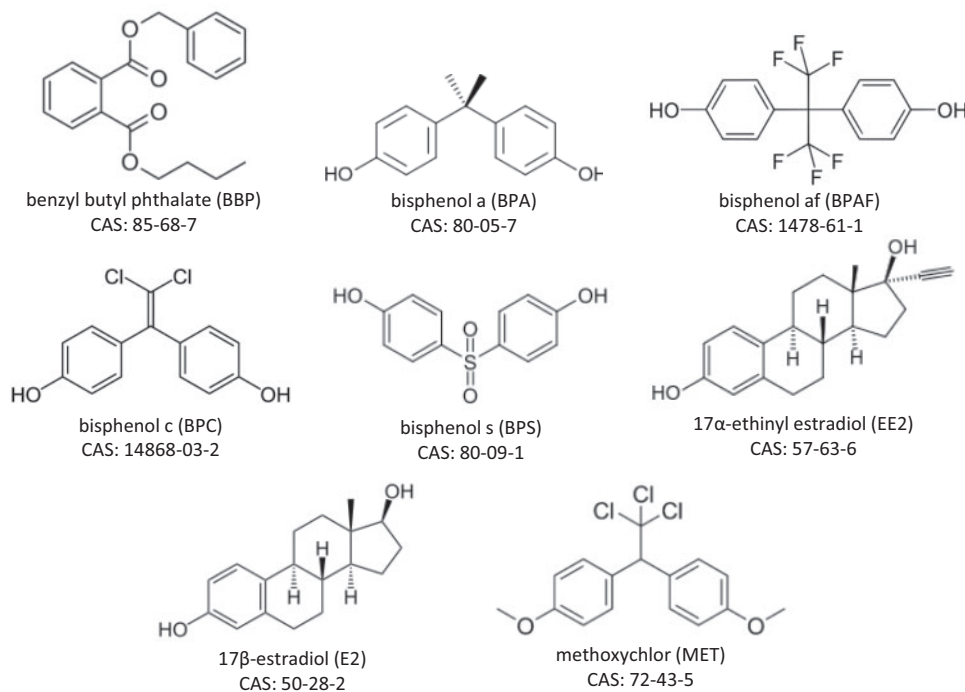


FIG. 1. Names, acronyms, structures, and Chemical Abstracts Service (CAS) numbers of test chemicals.

relative lack of data on BPAF and BPS, which are being used as replacements for BPA, we also wanted to compare the *in vivo* potencies of these to BPA. BPC was of interest herein because publications indicate that BPC is a potent ER agonist *in vitro*, however, to the best of our knowledge, there are no published *in vivo* data on this chemical.

The current study also included several mixture experiments with the above estrogenic chemicals to determine if the individual chemical *in vitro* and/or *in vivo* data could be used to predict the cumulative effects of these estrogens on uterine weight. Humans are potentially exposed to mixtures of estrogenic chemicals from several sources including endogenous synthesis of steroidal estrogens, dietary phytoestrogens, pharmaceuticals and personal care products, toxic substances and pesticides with estrogenic activity, and steroidal estrogens in source waters. In addition, fish and wildlife may be exposed to a range of natural estrogens and xenoestrogens in contaminated aquatic and terrestrial environments. Clearly, exposure to more than one estrogenic compound at a time is the rule rather than the exception (Bermudez et al., 2010; Leet et al., 2011). Component-based approaches to estimating the total potency of an estrogenic mixture involves using dose-response data on the individual compounds within the mixture to predict the overall mixture response (Rider and LeBlanc, 2005).

In the current study, we evaluated estrogenic activity *in vitro* using the T47D-KBluc estrogen receptor alpha (ER α)-mediated transcriptional activation reporter assay and *in vivo* using the adult ovariectomized rat uterotrophic assay (oral dosing), which is also largely ER α -mediated (Frasor et al., 2003). Overall, this study was designed to test the hypothesis that *in vitro* estrogen receptor activation assay data is predictive of the *in vivo* uterotrophic assay dose response to estrogens individually and as mixtures. Additionally, by testing this hypothesis, we intended to generate a dataset of *in vitro* and oral *in vivo* ER response data for chemicals spanning a wide range of potencies and ADME parameters for future physiologically based pharmacokinetic (PBPK) or reverse toxicokinetic (rTK) modeling efforts. The results of this study provide a demonstration of the complexity involved in extrapolating from *in vitro* to *in vivo* for orally administered endocrine active compounds.

MATERIALS AND METHODS

Test chemicals

All *in vitro* test chemicals were dissolved in either dimethyl sulfoxide (DMSO; Sigma Aldrich, St. Louis, Missouri) or ethanol (EtOH, 100%; Pharmco-AAPER, Brookfield, Connecticut) and all *in vivo* test chemicals were dissolved or suspended in corn oil (density 0.9 g ml⁻¹; Sigma Aldrich). BPA (CAS 80-05-7, lot 03105ES, cat no. 239658, purity = 99%), BPAF (CAS 1478-61-1, lot 12710JB, cat no. 257591, purity = 97%), BBP (CAS 85-68-7, lot 03405JH, purity = 99%, density 1.1 g ml⁻¹), E2 (CAS 50-28-2, lot 28H0818, cat no. E8875, purity = 98%), EE2 (CAS 57-63-6, lot 071M1492V, cat no. E4876, purity \geq 98%), and β -estradiol 3-benzoate (EB; CAS 50-50-0, lot 550-0025, cat no. e-9000, purity = 98%) were purchased from Sigma Aldrich. BPC (CAS 14868-03-2, lot WH66A-LQ, cat no. D3223, purity = 98%) was purchased from TCI America (Cambridge, Massachusetts). BPS (CAS 80-09-1, lot IF20140175, cat no. HS30-101, purity > 99%) was purchased from Ivy Fine Chemicals (Cherry Hill, New Jersey). MET (CAS 72-43-5, lot LB88314V, cat no. 4-9054, purity = 99.9%) was purchased from Supelco Analytical (St. Louis, Missouri). The ER antagonist ICI 182,780 (ICI; CAS 129453-61-8, batch 20A/116982,

purity 99%) was purchased from ICI Pharmaceuticals (United Kingdom).

T47D-KBluc transcriptional activation assay

In vitro estrogenic activity was assessed using the T47D-KBluc estrogen receptor transcriptional activation (ERTA) assay, which is responsive to estrogenic compounds spanning a broad range of potencies (Bermudez et al., 2012; Wilson et al., 2004). This assay was previously evaluated by a contract laboratory through a validation set of chemicals containing over 40 *in vitro* ER agonist reference chemicals that were selected from the Interagency Coordinating Committee on the Validation of Alternative Methods (ICCVAM) list or which had well characterized ER activity (data not published). T47D-KBluc cells have been stably transfected with a triplet estrogen-responsive element promoter-luciferase reporter gene construct. These cells originally expressed both endogenous ER α and ER β (Wilson et al., 2004), however we recently conducted gene expression array experiments, that include ER α and ER β , and transactivation experiments with ER α and ER β specific ligands with the cell culture maintained in our laboratory and determined that our cells now show little to no evidence of ER β expression and are dominated by ER α (Supplementary Table 1 and Supplementary Figure 1). Cell maintenance and assay execution were similar to the methods detailed in Wilson et al. (2004) and Bermudez et al. (2010). Briefly, T47D-KBluc (ATCC# CRL-2865) cell cultures were maintained at 37 °C and 5% CO₂ in vented 75 cm² culture flasks (Corning 430641) containing RPMI 1640 growth media (Gibco 13200-076; phenol red-free) supplemented with 10% fetal bovine serum (FBS; Hyclone #SH30071.03), 2 mM glutamine, 100 U ml⁻¹ penicillin, 100 μ g ml⁻¹ streptomycin, and 0.25 μ g ml⁻¹ amphotericin. All plastic culture materials including flasks, conical tubes, assay plates, and pipettes have been selected for those made of materials that minimize estrogenic contamination. Maintenance cells were passed into fresh media every 7 days followed by media renewal 3–4 days after subculture.

T47D-KBluc cell passages 45–115 were used to run the ERTA assays with cells that were withdrawn from culture media for 1 week prior to assay to remove all trace/residual estrogens. Withdrawal media consisted of RPMI supplemented with 10% dextran-coated charcoal-stripped FBS (DCC-FBS; Hyclone #SH30068.03) and was renewed once after 3–4 days and prior to running the assay. Cells were seeded at 10⁴ cells/100 μ l well⁻¹ in a 96-well luminometer plate (Costar 3610) in 5% DCC-FBS and incubated overnight. Chemical dosing occurred the following day (18–24 h after seeding plates).

Dosing of the attached cells involved preparation of chemical stock solutions in 100% DMSO or 100% EtOH. Stock solutions were prepared and stored at 4 °C in glass amber vials with Teflon-lined caps. Prior to exposure, stock solutions were diluted in 5% DCC-FBS to prepare the dosing solutions, with the maximum vehicle concentration not exceeding 0.1% (by volume). The withdrawal media on the cells was replaced with fresh dosing solution and cells were incubated overnight. Following 24 h incubation, cell viability was visually scored using a rating scale based on morphological cellular changes (eg, score 0 = non-cytotoxic/normal morphology and density, up to score 4 = severely cytotoxic/no visible cells) similar to the approach reported by Bhatia and Yetter (2008). Cells were then washed and lysed and luciferase activity was measured as relative light units (RLU), using a BMG LUMI-star 96-well plate luminometer (BMG Labtech; Durham, North Carolina).

All compounds were run on individual plates over a range of doses separated at half-log intervals. Each plate included 4 technical replicates (ie, wells) per dose group and each chemical was tested across ≥ 6 biological replicates (ie, unique cell passages). Each plate contained a vehicle blank, E2 standard curve (300 fM, 1 pM, 3 pM, 10 pM, and 30 pM), a dilution series of an individual test compound, antagonist (1 nM ICI) plus E2 (30 pM), and antagonist (1 nM ICI) plus test compound (concentration producing maximum induction). Cell performance in the assay was assessed by monitoring maximum fold induction above vehicle control (≥ 5 -fold at saturating E2 concentrations) and consistency of the E2 standard curve (ie, no significant change in EC_{50} over time) that was run on each replicate plate at the same dose range.

Uterotrophic assay

Ovariectomized Sprague Dawley rats of approximately 60 days of age were purchased from either Harlan Laboratories (Indianapolis, Indiana) or Charles River Laboratories (Raleigh, North Carolina). Rats were housed in pairs in clear, polycarbonate cages ($20 \times 25 \times 47$ cm) containing laboratory grade pine shavings as bedding and provided NTP 2000 diet (National Toxicology Program, NIEHS; Research Triangle Park, North Carolina) and filtered ($5 \mu\text{m}$) municipal tap water, *ad libitum*. Cages were kept in an animal facility room at the USEPA (Research Triangle Park, North Carolina), maintained on a 12:12 light:dark photoperiod (lights off at 18:00) at $20\text{--}22^\circ\text{C}$ and 45–55% relative humidity. Following arrival at the facility, rats were held for three weeks to allow time for uterine regression and for the animals to acclimate. These studies were conducted under protocols approved by the National Health and Environmental Effects Research Laboratory Institutional Animal Care and Use Committee at a facility accredited by the Association for Assessment and Accreditation of Laboratory Animal Care.

Rats were randomly assigned to treatment and given a single daily dose of test chemical, or chemical mixture, by oral gavage at a dose volume of $2.5 \text{ ml kg body weight}^{-1}$ for 4 consecutive days. Overt toxicity was defined as signs of diarrhea, malaise, and/or general lack of grooming. Dosing was discontinued immediately in the presence of any of these signs and uterine weights of these rats were excluded from the analyses. All compounds (BPA, BPAF, BPC, BPS, BBP, E2, EE2, and MET) were tested individually; dose groups and sample sizes for each compound are listed in [Supplementary Table 1](#).

In each experimental block, there was a negative control group (corn oil only), a positive control group (subcutaneous injection of $1 \mu\text{g}$ EB), and 3–8 dose levels of an individual test chemical or mixture, each consisting of 3–6 rats. Rats were euthanized by decapitation 3 h after the final dose, trunk blood was collected and uterine tissue was excised, trimmed of fat tissue, and weighed with luminal fluid for a “wet weight”. Punctured uterine tissue was then blotted on absorbent paper to drain fluid and the uterus was reweighed for a “blotted weight”. “Fluid weight” was calculated as the mass difference between the uterine wet weight and blotted weight.

Uteri were preserved in Bouin’s solution, transferred to 100% EtOH after 24 h, and sent to Experimental Pathology Laboratories, Inc. (EPL; Durham, North Carolina) for histopathological examination. Tissues were imbedded in paraffin, sectioned at $5 \mu\text{m}$, stained with hematoxylin and eosin (H&E), and examined by an EPL pathologist certified by the American Board of Pathology.

Serum was isolated from trunk blood via centrifugation ($10\,000 \times g$; 15 min) using Becton Dickinson (Franklin Lakes,

New Jersey) vacutainer tubes and stored in 1.5 ml siliconized microcentrifuge tubes at -80°C for future chemical analyses. In-house analyses for this study were limited to E2 quantitation using radioimmunoassay (estradiol ultra-sensitive RIA; Beckman Coulter, Indianapolis, Indiana; item #DSL-4800; assay range $5\text{--}650 \text{ pg ml}^{-1}$) according to manufacturer specifications.

In vivo binary and multi-chemical mixture studies were performed using fixed ratio dilutions, where the ratio of the individual components making up the total dose was set with the target that each individual compound contributed equally to the response. The ratios were determined based on the potency factor of each compound derived from individual dose response curves for increased blotted uterine weight in the uterotrophic assay. Binary mixtures of BPAF:MET were set at a ratio of 1.5:1 and BPS:MET were set at 2:1 (total dose and sample size reported in [Supplementary Table 2](#)). For the multi-chemical mixture, the top dose was represented by each compound (BPAF, BPC, BPS, EE2, and MET) at their respective individual ED_{50} values for blotted uterine weight.

Calculations and statistical analyses

Data analyses were performed using GraphPad Prism (v5.0, San Diego, California) and SAS (Cary, North Carolina). Briefly, raw data from the T47D-KBluc ERTA were normalized to background, \log_{10} transformed, and converted to percentage maximum response based on saturating levels of E2. Specifically, luminescence (RLU) for each well was normalized to the mean vehicle control (ie, fold induction = well RLU/mean vehicle RLU) and then \log_{10} transformed to correct for heterogeneity of variance (\log fold induction). E2 standard concentrations (molar; x-axis) were then \log_{10} transformed and plotted against the \log fold induction (y-axis) and fit with a 4-parameter logistic regression (bottom of curve constrained to 0) to determine the maximum E2 \log fold induction (ie, top of regression curve). \log fold induction values for each sample well were then converted to percentage of maximum E2 response (%E2 max) by dividing by the top parameter of the E2 standard curve and multiplying by 100. Sample concentrations were then \log transformed and plotted as a function of %E2 max values and fit using 4-parameter logistic regression (bottom = 0) for determination of slope and EC_{50} . Finally, relative potency factors (RPF) were calculated for each compound using EE2 as the reference chemical ($\text{RPF} = \text{EE2 } EC_{50} / \text{compound } EC_{50}$).

A similar approach was taken for analyses of *in vivo* individual compound and mixture uterotrophic assay data. For each of the three uterine endpoints (wet weight, blotted weight, and fluid weight), within a given test block of animals, the weights were \log_{10} transformed, then blank corrected by subtracting the mean of the negative control weights, and finally converted to percentage of maximum response (% max) by dividing by the mean weight of the positive controls and multiplying by 100. Dose response curves were plotted using 4 parameter logistic regression with % max as a function of the \log_{10} transformed oral gavage dose ($\text{mg kg}^{-1} \text{ day}^{-1}$) and curve top and bottom constrained to 100 and 0, respectively. RPFs were again calculated using EE2 as the reference compound. Additionally, analysis of variance (ANOVA) with Dunnett’s pairwise comparison was performed on \log_{10} transformed weights using SAS (proc GLM) to identify statistically significant increases as compared to the negative controls for determination of no observed effect level (NOEL) and lowest observed effect level (LOEL) ([Supplementary Tables 1 and 2](#)).

In vivo uterotrophic response predictions for individual compounds were determined by shifting the EE2 reference compound curve along the x-axis based on the *in vitro* RPF.

For example, a test compound with an *in vitro* RPF of 0.005 would result in the EE2 curve being right shifted by a factor 200 (ie, $1/0.005 = 200$). This was performed separately for each of the 3 uterine endpoints (wet, blotted, and fluid) using the EE2 reference curve for the appropriate endpoint.

Mixture response predictions were performed using the dose addition (DA) model with both individual chemical *in vitro* and *in vivo* data, separately. DA model predictions for the mixture responses were calculated using the equation:

$$R = \frac{1}{1 + \frac{1}{\left(\sum_{i=1}^n \frac{c_i}{ED_{50i}}\right)^{\rho'}}$$

where R = mixture response, c_i = dose of chemical i , ED_{50i} = dose of chemical i resulting in 50% response, and ρ' = mean Hill slope of compounds in model. Predictions based on *in vitro* data were made using the predicted *in vivo* ED_{50} s derived from the *in vitro* RPFs, as described above, for the individual compounds in the DA model. The Hill slope used in the *in vitro* data DA model was the slope of the *in vivo* EE2 curve. Predictions based on *in vivo* data were made using the actual *in vivo* ED_{50} values and mean Hill slopes for component compounds. DA model accuracy was assessed using the 95% confidence intervals (CI) for the observed mixture data with the DA model predictions. The prediction was considered significantly different if it fell outside of the 95% CI of the observed data.

RESULTS

In Vitro Dose Response Assessment

All compounds (BBP, BPA, BPAF, BPC, BPS, E2, EE2, and MET) were positive for estrogenic activity in the T47D-KBluc assay (Figure 2a). In general, compounds displayed similar slopes except for BBP (4.07) and MET (2.49) which were slightly steeper than the remaining compounds (mean slope, 1.61 ± 0.22 ; Table 1). EC_{50} values spanned 6 orders of magnitude with EE2 the most potent (9.98×10^{-13} M) and MET the least potent (1.47×10^{-6} M) compounds. As expected, the environmental chemicals (BBP, BPA, BPAF, BPC, BPS, and MET) were

considerably less potent than E2 (RPF 0.594), or EE2 (RPF 1.00), and clustered toward the higher end of the x-axis, spanning ~ 3 orders of magnitude (RPF range, 0.000175 – 0.00000679). Interestingly, BPC (EC_{50} 5.71×10^{-9} M) was the most potent environmental estrogen that we have examined to date. Finally, the efficacies (maximum fold induction reached) of the compounds were variable with MET only reaching $\sim 70\%$ of the maximum fold induction potential, whereas BPC reached $\sim 115\%$ fold induction relative to the E2 standard.

General Toxicity In Vivo

Dosing was discontinued on the third or fourth day in the 500 (3 of 11 rats), 650 (2 of 3 rats), and 1000 (3 of 3 rats) $\text{mg kg}^{-1} \text{day}^{-1}$ dose groups for the BPAF + MET mixture due to signs of overt toxicity. No other chemical or mixture dosages caused general toxicity. For some of the individual compounds (E2, BPA, BPAF, and MET), including the positive control (EB), and all of the mixtures, there were mild, but significant reductions in weight change during dosing versus the negative control (range, 0.1–9.1% mean body weight reduction between days 1 and 4; Supplementary Tables 2 and 3). These decreases were not unexpected because it is well documented that estradiol suppresses feeding behavior, food intake, and reduces body weight gain (Bonavera et al., 1994). In no case did the weight loss exceed 15.5% of the initial body weight in any animal.

TABLE 1. *In vitro* (T47D-KBluc) dose response curve parameters (slope, EC_{50} (M)) and relative potency factors (RPF; using EE2 as reference compound) for eight estrogenic compounds

Chemicals	Slope \pm SE	EC_{50} (95% CI)	RPF
EE2	1.81 ± 0.19	9.98×10^{-13} (8.81×10^{-13} – 1.13×10^{-12})	1.00
E2	1.46 ± 0.09	1.68×10^{-12} (1.52×10^{-12} – 1.86×10^{-12})	0.594
BPC	1.46 ± 0.11	5.71×10^{-9} (5.03×10^{-9} – 6.48×10^{-9})	0.000175
BPAF	1.40 ± 0.11	2.06×10^{-8} (1.81×10^{-8} – 2.35×10^{-8})	0.0000485
BPA	1.61 ± 0.18	1.27×10^{-7} (1.08×10^{-7} – 1.50×10^{-7})	0.00000784
BPS	1.93 ± 0.27	6.43×10^{-7} (5.30×10^{-7} – 7.80×10^{-7})	0.00000155
BBP	4.07 ± 0.87	1.31×10^{-6} (1.12×10^{-6} – 1.53×10^{-6})	0.000000763
MET	2.49 ± 0.38	1.47×10^{-6} (1.26×10^{-6} – 1.72×10^{-6})	0.000000679

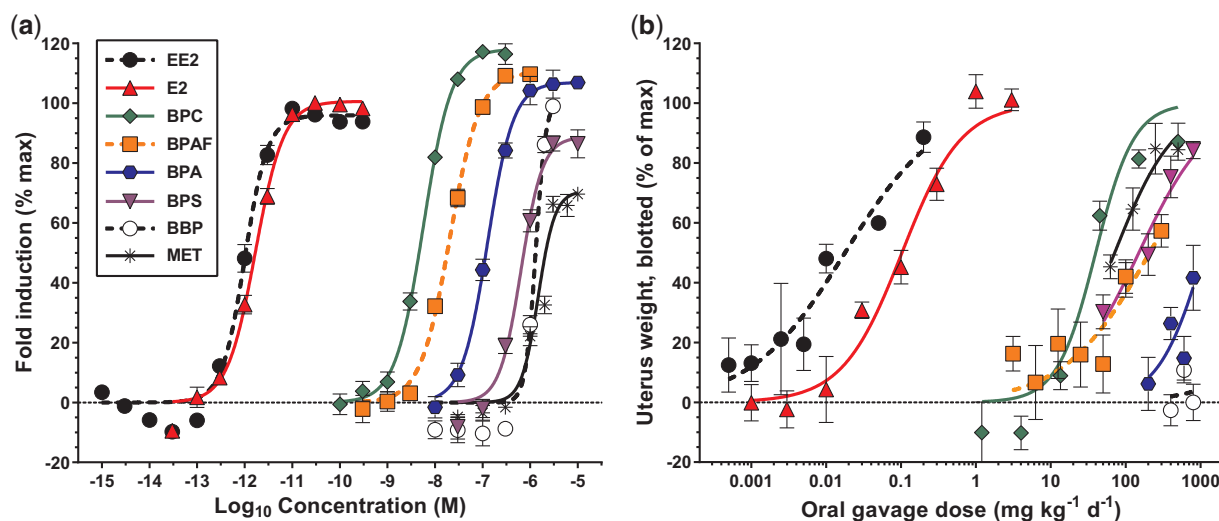


FIG. 2. Individual chemical *in vitro* (T47D-KBluc assay, panel a) and *in vivo* (ovariectomized rat uterotrophic, oral administration, panel b) dose response curves. Symbols and error bars represent mean \pm standard error, dashed horizontal line indicates zero baseline. Color version available in online version of manuscript.

In Vivo Individual Compound Dose Response Assessments

All compounds were positive for estrogenic activity in the ovariectomized rat uterotrophic assay using oral exposure except for BBP, which has been negative in other studies and was a false positive *in vitro* (Figure 2b and Supplementary Figure 2). All uteri with increased weight had a histopathological correlate of increased columnar cell height and differentiation (Supplementary Table 4). Individual dose response curve parameters were determined for wet, blotted, and fluid uterine weights (Table 2). Uterus blotted weight was the most sensitive (lowest ED₅₀) and least variable (19.9% average CV for tissue mass across treatments, versus 24.1% and 49.0% CV for wet and fluid masses, respectively) endpoint for all compounds. Further, similar LOELs and NOELs were derived using either wet or blotted uterus weight (Supplementary Table 2). As such, the remainder of this article will focus on the uterus blotted weight endpoint.

Similar to the T47D-KBluc assay results, uterotrophic ED₅₀ values spanned 5 orders of magnitude ranging from the most potent compound EE2 (0.0179 mg kg⁻¹ day⁻¹) to the least potent compound BPA (1118.0 mg kg⁻¹ day⁻¹). EE2 (RPF 1.00) and E2 (RPF 0.173) were again considerably more potent than the environmental chemicals (RPF range, 0.0000160–0.000449); however, BPC (39.8 mg kg⁻¹ day⁻¹) was the most potent xenoestrogen *in vivo* and among the most potent xenoestrogens tested to date. Interestingly, BPS was 5.1-fold less potent than BPA *in vitro*; however, *in vivo* BPS was 7.5-fold more potent than BPA. MET was the least potent compound *in vitro* but the second most potent compound *in vivo*.

Dose-related histological changes, including increased epithelial and glandular cell height, were seen in the uterine tissue of rats treated with BPAF (50 mg kg⁻¹ day⁻¹), EE2 (0.05 mg kg⁻¹ day⁻¹), MET (125 mg kg⁻¹ day⁻¹), BPC (1/6 females responded at 45 mg kg⁻¹ day⁻¹ and 6/6 at 150 mg kg⁻¹ day⁻¹), and BPS (200 mg kg⁻¹ day⁻¹), whereas BPA only induced estrogen-like changes in the uterus of 1/6 rats in the highest dose group (800 mg kg⁻¹ day⁻¹) and BBP had no effect on uterine histology (Supplementary Table 4). Tissues from animals treated with E2 were not examined. These effects are consistent with the RPF ranking based upon blotted uterine weights, however, statistically significant uterine weight increases were occasionally detected at dose levels that produced minimal/no histological effects. The RPF ranking based on estrogen-like histological changes in the uterus is EE2 > BPC ≈ BPAF > MET > BPS > BPA ≫ BBP.

Serum E2 concentrations were significantly elevated above corn oil controls in the 0.1, 0.3, 1, and 3 mg kg⁻¹ day⁻¹ dose groups (P < .05), corresponding to E2 concentrations of 16.3 ± 1.3, 22.1 ± 3.6, 165.8 ± 93.6, and 175.6 ± 17.3 pg ml⁻¹, respectively. E2 was 37.1-fold more potent at inducing ER-mediated transcriptional activation in the T47D-KBluc assay (EC₅₀ = 1.7e–12 M) than it was at inducing uterine weight gain in the uterotrophic assay (EC₅₀ = 6.3e–11 M; Figure 3), based on serum E2 concentration following oral administration.

Comparison of In Vitro to In Vivo Predictions With In Vivo Observed Responses

The predictive capacity of *in vitro* screening results for estimating *in vivo* chemical responses was one of the primary interests of this study. We utilized *in vitro* RPFs as correction factors applied to the EE2 reference compound *in vivo* dose response curve to estimate *in vivo* response curves for the remaining seven compounds (Figure 4 and Tables 2 and 3). This extrapolation resulted in just two compounds (BPA and BPAF) having

TABLE 2. *In vivo* (ovariectomized rat uterotrophic assay, oral gavage) dose response curve parameters [slope, ED₅₀ (mg kg⁻¹ day⁻¹)] and relative potency factors (RPF; using EE2 as reference compound) across 3 endpoints (uterus blotted weight, uterus wet weight, and uterine fluid weight) for 7 estrogenic compounds.

Chemicals	Blotted			Wet			Fluid		
	Slope ± SE	ED ₅₀ (95% CI)	RPF	Slope ± SE	ED ₅₀ (95% CI)	RPF	Slope ± SE	ED ₅₀ (95% CI)	RPF
EE2	0.69 ± 0.14	0.0179 (0.00922–0.0346)	1.00	1.08 ± 0.17	0.0208 (0.0138–0.0315)	1.00	1.89 ± 0.56	0.0216 (0.0123–0.0379)	1.00
E2	1.09 ± 0.17	0.103 (0.0756–0.138)	0.173	1.17 ± 0.18	0.166 (0.127–0.216)	0.126	1.48 ± 0.29	0.249 (0.190–0.328)	0.0866
BPC	1.70 ± 0.34	39.8 (30.6–51.7)	0.000449	0.92 ± 0.15	95.4 (66.1–137.7)	0.000218	0.85 ± 0.25	286.8 (148.7–553.3)	0.0000753
BPAF	0.75 ± 0.19	209.1 (111.0–394.0)	0.0000854	0.62 ± 0.15	465.2 (198.0–1093.0)	0.0000448	0.47 ± 0.26	872.9 (49.7–15346.0)	0.0000247
BPA	1.40 ± 0.70	1118.0 (529.1–2360.0)	0.0000160	1.34 ± 0.72	1286.0 (504.8–3277.0)	0.0000162	1.26 ± 1.13	1737.0 (225.0–13418.0)	0.0000124
BPS	0.91 ± 0.14	150.0 (113.8–197.5)	0.000119	0.83 ± 0.13	259.2 (194.2–345.9)	0.0000804	1.05 ± 0.28	676.1 (412.7–1108.0)	0.0000319
MET	1.08 ± 0.22	71.4 (49.6–102.8)	0.000250	0.90 ± 0.13	126.7 (98.9–162.3)	0.000164	0.83 ± 0.20	104.4 (67.9–160.5)	0.000207

No values reported for BBP because it was negative in the oral uterotrophic assay.

predicted ED₅₀ values within the 95% CI of the observed ED₅₀. BPC was the only chemical that gave variable prediction accuracies; the predicted wet uterus weight ED₅₀ fell within the 95% CI of the observed data whereas predicted blotted and fluid weights did not. The remaining four chemicals had predicted ED₅₀ values outside of the 95% CI of the observed data for all three endpoints. E2 was 3.4-fold less potent than predicted for blotted uterus weight and BBP was a false positive based on *in vitro* results. Conversely, BPS and MET had predicted-to-observed ratios (P:O) of 76.6- and 368.3-fold, respectively, indicating that the potencies were considerably underestimated by *in vitro* predictions.

In Vivo Mixture Dose Response Assessments

The binary and multi-chemical mixtures were potent stimulators of the uterotrophic response. All mixtures conformed to the dose addition model with individual chemical *in vivo* dose response data producing more accurate ED₅₀ predictions than the *in vitro* data (Table 4 and Figure 5). For both the BPAF + MET and BPS + MET binary mixtures, the predicted ED₅₀ values for blotted uterine weight fell outside of the 95% CI for the observed values but only by a narrow margin (P:O of 1.5 and 0.7 for BPAF + MET and BPS + MET, respectively). In contrast, the predicted ED₅₀ values using *in vitro* data were 7.7- and 86.8-fold greater than observed for BPAF + MET and BPS + MET, respectively. For the multi-chemical mixture, the predicted ED₅₀ using *in vivo* data (93.2 mg kg⁻¹ day⁻¹) was nearly identical to the observed ED₅₀ (95.5 mg kg⁻¹ day⁻¹), as opposed to the predicted ED₅₀ using *in vitro* data, which was 2.5-fold greater (237.3 mg kg⁻¹ day⁻¹) and outside of the observed 95% CI. For uterus wet weight, all predicted ED₅₀ values using *in vivo* data were within the observed 95% CI, whereas the *in vitro* predictions were 2.0- to 85.9-fold greater. Further, for uterine fluid weight, the predicted ED₅₀ values using *in vivo* data for the multi-chemical mixture and BPAF + MET fell within the observed 95% CIs, whereas the BPS + MET prediction was just below the 95% CI (P:O of 0.7). All of the uterine fluid weight ED₅₀ predictions using *in vitro* data were outside of the observed 95% CIs, but to a lesser degree (P:O range of 1.4- to 50.5-fold) than the wet and blotted uterine weight predictions.

DISCUSSION

The “Toxicity Testing in the 21st Century” (Tox21) vision and strategy was developed by a National Academy of Sciences committee at the request of the USEPA to address the shortcomings of current regulatory toxicity testing, which relies heavily on the use of whole animals. One goal of the Tox21 strategy is to reconfigure toxicity testing using high- and medium-throughput *in vitro* screening assays (NAS, 2007). Nevertheless, the committee acknowledged that a major challenge faced in reaching this goal will be accounting for whole animal toxicokinetics, including absorption, metabolism, distribution, and excretion (ADME). Here, we demonstrate some of the challenges associated with the use of an *in vitro* estrogen screening system for extrapolating to an *in vivo* estrogenic response. Specifically, the uterotrophic responses for BPA, BPAF, and BPC were accurately predicted by an *in vitro* assay, whereas the potencies of BPS and MET were under-predicted, the potency of E2 was over-predicted, BBP was a false positive, and there were no false negatives in this group of test chemicals. Further, the cumulative response to mixtures of xenoestrogens *in vivo* was better predicted using individual compound *in vivo* dose response data than using corresponding *in vitro* data. The deficiencies in

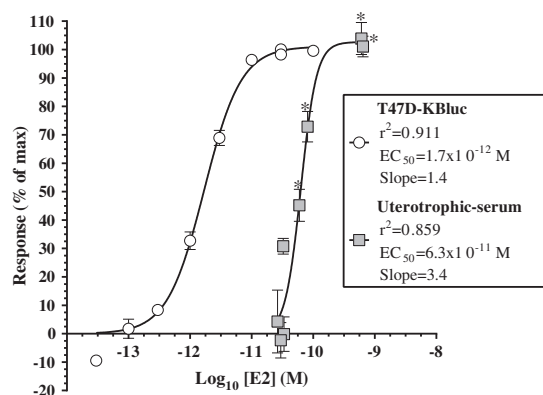


FIG. 3. Comparison of dose response curves for 17 β -estradiol (E2) in the T47D-KBluc assay (white circles) and oral uterotrophic assay (grey squares). Uterotrophic assay response data presented as a function of serum E2 concentration determined by radioimmunoassay (oral dose presented in prior figures). Treatment groups with statistically significant increases in serum E2 concentration as compared to corn oil control identified by asterisks ($P < .05$). Data points represent mean \pm standard error.

predicting *in vivo* response from *in vitro* screening data can primarily be explained by the metabolic activation or inactivation of the xenoestrogens used in this study.

Benzyl-butyl phthalate (BBP) is a plasticizer used primarily in polyvinylchloride products and was a false positive (ie, active *in vitro* but inactive *in vivo*) in the present study, which is consistent with the literature. Studies have reported on the *in vitro* estrogenicity of BBP including binding to the estrogen receptor and activating transcription of downstream ER genes (Blair *et al.*, 2000; Jobling *et al.*, 1995; Picard *et al.*, 2001). Conversely, *in vivo* evaluations of BBP have consistently demonstrated no estrogen-like stimulation in the uterotrophic assay with oral dosing up to 2240 mg kg⁻¹ day⁻¹ (Brady *et al.*, 2000; Zacharewski *et al.*, 1998) or subcutaneous dosing up to 5000 mg kg⁻¹ day⁻¹ (Brady *et al.*, 2000). Further, at doses up to 200 mg kg⁻¹ day⁻¹ BBP has no effect on vaginal opening in the female pubertal assay (Ahmad *et al.*, 2015). This lack of estrogenic response *in vivo* is because BBP is rapidly hydrolyzed to non-estrogenic monoester metabolites following both oral and subcutaneous administration and the parent material never appears in the serum following dosing (Brady *et al.*, 2000).

17 β -Estradiol (E2) was positive in both assays but the oral *in vivo* potency of E2 was over-predicted using *in vitro* assay data. E2 is the predominant endogenous estrogen in vertebrate endocrine systems, is naturally excreted into wastewater, and has been detected in sewage treatment plant effluents, surface water, and drinking water treatment plant source water (Richardson and Ternes, 2005). To the best of our knowledge, there are no studies reporting detectable concentrations of E2 in treated drinking water; however, due to its presence in surface water and endocrine disrupting potential, E2 was included in the USEPA Contaminant Candidate List 3 (CCL3) (USEPA, 2009) indicating the need to further study the response to oral exposure of this compound. Orally administered E2 is rapidly converted to the less potent estrogens, estrone and estrone sulfate, via first-pass hepatic metabolism (Heller, 1940; Powers *et al.*, 1985). The rapid conversion leads to little E2 being absorbed into systemic circulation and therefore a less potent response than would be predicted from *in vitro* data. Further, the proportion of orally absorbed E2 that does make it into systemic circulation is extensively bound by plasma proteins (~95%) and unavailable for tissue uptake (Plowchalk and

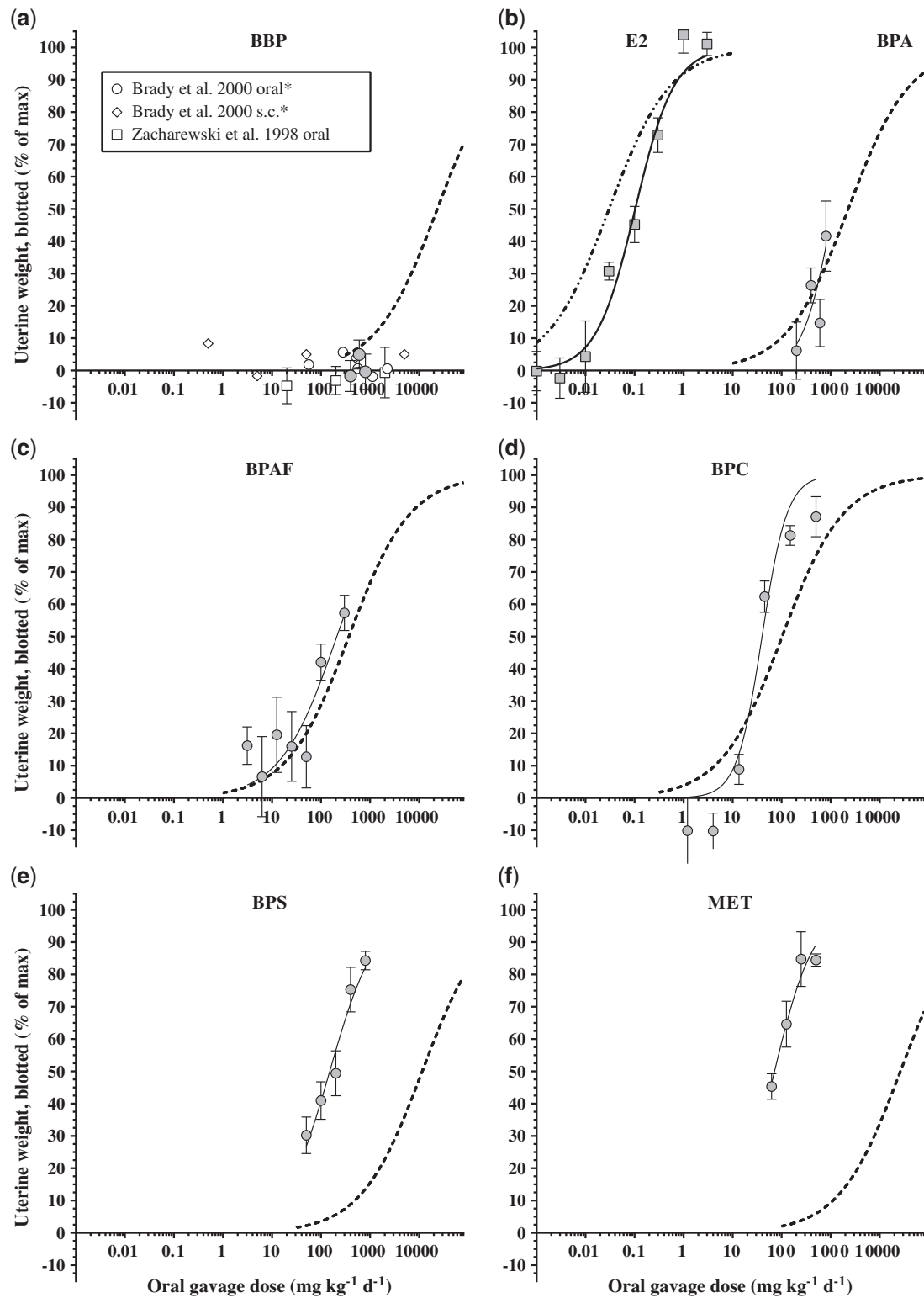


FIG. 4. Predicted (black dashed lines) versus observed (grey circles and black solid lines (except panel b where E2 is represented by grey squares) dose response curves for uterus blotted weight in the ovariectomized rat uterotrophic assay (oral gavage). Predicted lines were generated using *in vitro* (T47D-KBluc) relative potency factors with 17 α -ethinyl estradiol as a reference compound. Panel (a) includes additional data estimated from prior studies on the uterotrophic response to BBP [“data from Brady et al. (2000) represents uterine wet weight”].

TABLE 3. Predicted ED₅₀ (mg kg⁻¹ day⁻¹) values for individual chemicals in the ovariectomized rat uterotrophic assay across three endpoints (uterus blotted weight, uterus wet weight, and uterine fluid weight).

Chemicals	Blotted		Wet		Fluid	
	Predicted ED ₅₀	P:O	Predicted ED ₅₀	P:O	Predicted ED ₅₀	P:O
EE2	–	–	–	–	–	–
E2	0.030	0.3	0.035	0.2	0.036	0.1
BPC	102.1	2.6	119.1	1.2	123.5	0.4
BPAF	368.3	1.8	429.8	0.9	445.5	0.5
BPA	2276.2	2.0	2656.2	2.1	2753.1	1.6
BPS	11497.1	76.6	13416.5	51.8	13906.0	20.6
BBP	23387.6	–	27292.0	–	28287.8	–
MET	26284.2	368.3	30672.2	242.1	31791.3	304.5

Predictions were calculated using *in vitro* (T47D-KBluc) relative potency factors with EE2 as the reference compound. Predicted to observed ratios (P:O) indicate accuracy of prediction (P:O = 1.0 indicates perfect prediction).

TABLE 4. Observed and predicted (Pred.) ED₅₀ values (mg kg⁻¹ day⁻¹) for binary (BPAF + MET, BPS + MET) and multi-chemical (BPAF + BPC + BPS + EE2 + MET) mixtures in the ovariectomized rat uterotrophic assay across 3 endpoints (blotted uterus weight, wet uterus weight, uterine fluid weight).

Mix.	Blotted				Wet				Fluid						
	Observed ED ₅₀ (95% CI)	<i>In vivo</i>		<i>In vitro</i>		Observed ED ₅₀ (95% CI)	<i>In vivo</i>		<i>In vitro</i>		Observed ED ₅₀ (95% CI)	<i>In vivo</i>		<i>In vitro</i>	
		Pred.	P:O	Pred.	P:O		Pred.	P:O	Pred.	P:O		Pred.	P:O	Pred.	P:O
BPAF+MET	79.3 (54.7–115.0)	118.0	1.5	709.8	7.7	197.4 (172.8–225.5)	224.9	1.1	608.2	3.6	249.4 (213.0–291.9)	221.3	0.9	735.6	2.9
BPS+ MET	162.6 (128.5–205.7)	110.1	0.7	14108.0	86.8	191.6 (151.0–243.1)	192.9	1.0	16463.0	85.9	337.6 (243.8–467.5)	241.3	0.7	17064.0	50.5
BPAF+ BPC+ BPS+ EE2+ MET	95.5 (72.1–126.4)	93.2	1.0	237.3	2.5	135.7 (104.0–176.9)	164.1	1.2	278.0	2.0	211.5 (161.0–277.9)	222.8	1.1	288.2	1.4

Predicted values were calculated using individual chemical response data from *in vivo* (uterotrophic) and *in vitro* (T47D-KBluc) experiments and the dose addition model. Predicted to observed ratios (P:O) indicate the accuracy of the prediction (P:O = 1 represents perfect prediction).

Teeguarden, 2002). The serum concentrations reported here represent total E2 concentration (ie, free and albumin bound compound), which likely explains the 37-fold difference in biological response between the T47D-KBluc and uterotrophic (as a function of serum [E2]) assays for E2.

In contrast to BBP and E2, MET is a strong example of a compound that undergoes *in vivo* metabolic activation following ingestion. MET is a pesticide that was registered in 1948 as a replacement for DDT but was cancelled for use in the U.S. in 2003 after being listed as persistent, bioaccumulative, and toxic (PBT) by the USEPA Toxics Release Inventory (TRI) program (USEPA, 2004). MET is considered a proestrogen because the parent compound is a less potent estrogen than several of its metabolites, including the demethylated metabolite, 2,2-bis-(*p*-hydroxyphenyl)-1,1,1-trichloroethane (HPTE) (Cummings, 1997; Hu and Kupfer, 2002). MET is not unique in terms of compounds known to be activated *in vivo*. For example, certain phytoestrogens, including 4'-methoxy-coumestrol and formononetin, are enhanced through demethylation to more potent estrogenic compounds, especially in ruminants, where they are responsible for "clover disease" in sheep (Adams, 1995; Cox and Braden, 1974). The potential to underestimate the potency of proestrogens, like MET, in the absence of *in vivo* activating enzymes is a primary concern when considering replacement of *in vivo* assays with *in vitro* assays. The underestimation of potency using an *in vitro* screen carries the risk of falsely classifying the compound as negative for activity and therefore precluding it from further testing. On the other hand, overestimation of

chemical potency using an *in vitro* screening assay may result in increased cost and time of further testing.

Similar to MET, although less extensively studied, BPS displayed considerably higher potency *in vivo* than *in vitro*. BPS has recently received increasing attention due to its use as a replacement for BPA in consumer products including plastics and thermal papers (Rochester and Bolden, 2015). BPS has been detected in paper currency, paper products, foodstuffs, and urine at concentrations similar to those previously reported for BPA (Liao and Kannan, 2013; Liao et al., 2012a, b). The ability of BPS to activate ER *in vitro* has been documented (Molina-Molina et al., 2013; Rosenmai et al., 2014), but there is a general lack of *in vivo* studies for comparison. Kang et al. (2014) and Hashimoto et al. (2001) indicated that BPS was biotransformed to more potent metabolites after incubation with rat liver S9; whereas, Le Fol et al. (2015) performed *in vitro* biotransformation studies and reported that the conjugates of BPS did not exhibit ER activity. Yamasaki et al. (2004) determined that BPS was positive in the immature female rat uterotrophic assay using subcutaneous injection, however the responses were not dose dependent (ie, increased uterine weight at 20 and 500, but not 100 mg kg⁻¹). The results of the current study clearly indicate that BPS produces dose-related increases in uterine weight following oral administration and that BPS is 7.5-fold more potent than BPA. Further studies of BPS would need to be conducted to determine how the multigenerational effects of BPS compare with the potential human and environmental exposure levels.

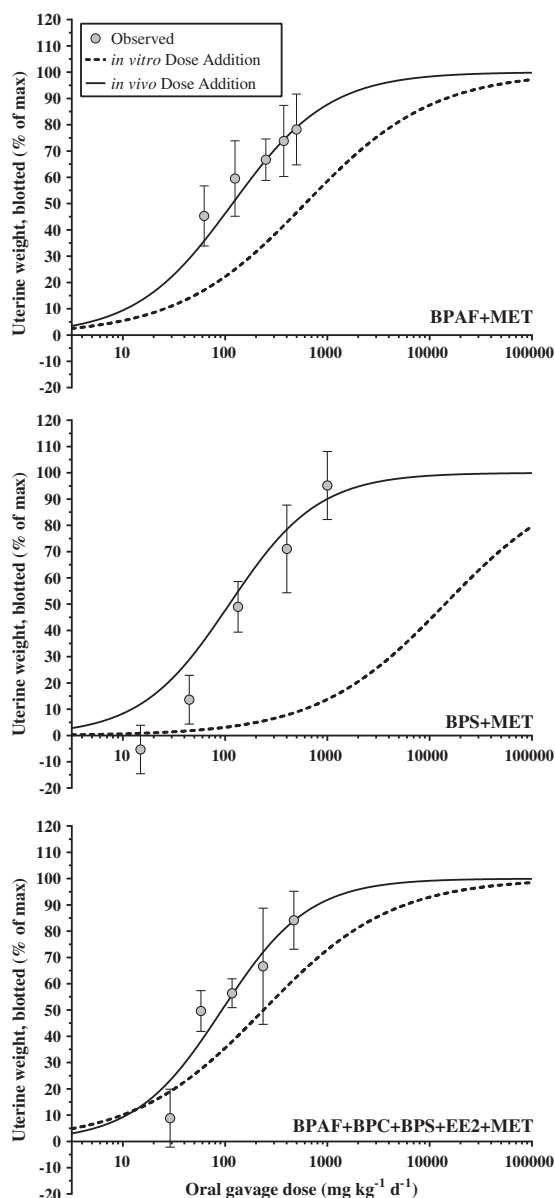


FIG. 5. Observed data (gray circles) versus dose addition (DA) models of binary (BPAF + MET; BPS + MET) and multi-chemical (BPAF + BPC + BPS + EE2 + MET) mixtures in the ovariectomized rat uterotrophic assay (oral gavage). DA models were generated using individual compound *in vitro* (T74D-Kbluc, black dashed lines) and *in vivo* (black solid lines) response data. Observed data represent mean \pm 95% CI; statistical significance based on model predictions overlapping the observed 95% CI.

Bisphenol-A (BPA) is a widely used additive in plastics and resins and is commonly recognized for its estrogenic properties and ubiquitous exposure to humans and wildlife (Fan et al., 2013; Liao and Kannan, 2013). Despite widespread exposure to BPA, toxicokinetic evaluations performed in rodents, humans, and non-human primates have indicated that the estrogenic parent compound does not accumulate in the body, but rather undergoes rapid and highly efficient metabolism, primarily glucuronidation and sulfation, and excretion following oral administration (Draganov et al., 2015; Kurebayashi et al., 2002; Pottenger et al., 2000; Teeguarden et al., 2011; Tominaga et al., 2006; Volkel et al., 2002). Although BPA reproducibly produces estrogenic responses *in vitro*, adverse *in vivo* effects associated

with exposure to BPA are highly debated (Sharpe, 2010). Herein, BPA was active *in vivo* and the potency was accurately predicted from the *in vitro* ERTA screening data; however, it was the least potent oral estrogen tested aside from BBP, which was negative *in vivo*.

The *in vivo* potency of BPAF was also accurately predicted from the *in vitro* ERTA screening data. BPAF is a fluorinated analogue of BPA used in the production of fluoro-elastomers and polycarbonates for food processing applications (NTP, 2008). There is currently very little information available on exposure levels of this compound; however, it has been detected in Chinese surface and drinking water (Song et al., 2012) and indoor dust in Korea (Liao et al., 2012c). BPAF has been shown to bind ER α (Laws et al., 2006; Matsushima et al., 2010), activate ER regulated gene transcription (Bermudez et al., 2010), stimulate MCF-7 cell proliferation (Hashimoto et al., 2001; Perez et al., 1998), and stimulate a uterotrophic response via subcutaneous administration (Akahori et al., 2008; Yamasaki et al., 2003). Toxicokinetic studies of BPAF following oral administration indicated that it is largely glucuronidated and/or sulfated and excreted in the feces similar to BPA, although more slowly (Waidyanatha et al., 2015; Yang et al., 2012). Currently, BPAF is undergoing extensive evaluations for *in vivo* toxicity by the National Toxicology Program.

Bisphenol-C (BPC) is also a halogenated analogue of BPA that is used in polymers and plastics to impart thermal stability and fire-resistance. However, BPC has undergone considerably less toxicity evaluation than BPAF and there is a paucity of information on human exposure. Similar to our current observations, *in vitro* studies on the ER binding affinity (Blair et al., 2000) and ER transcriptional activation and cell proliferation (Delfosse et al., 2012) activities of BPC have reported this compound to be more potent than other BPA analogues, as well as most other environmental chemicals. To the best of our knowledge, the current study represents the only available *in vivo* data on the estrogenic activity of BPC. Indeed, BPC is one of the most potent environmental estrogens ever tested by our laboratory both *in vitro* and *in vivo*. It is interesting to note that the chemical structure of BPC is nearly identical to HPTE, the potent estrogenic metabolite of MET [HPTE elicits a \sim 100-fold more potent uterotrophic response than MET using *i.p.* administration (Bulger et al., 1978)], differing only by a single chlorine atom. As such, further *in vivo* study of BPC is warranted given the high potency (28-fold more potent than BPA) and potential for exposure.

It has become well established that the dose response of mixtures of endocrine disruptors exhibiting a similar *in vitro* or *in vivo* effect (eg, estrogenic, androgenic, antiandrogenic) can be accurately predicted using dose addition modeling (Kortenkamp, 2007). Understanding joint action of compounds is important in trying to identify the most protective risk assessment approach. Compounds that conform to dose additive toxicity make good candidates for cumulative assessment. The limiting factor is the suitability and availability of the individual chemical data used in the modeling effort (Tinwell and Ashby, 2004). Here, mixture predictions of the *in vivo* uterotrophic effect generated using individual chemical *in vivo* data were found to be considerably more accurate when compared with the mixture models generated using individual chemical *in vitro* data. As previously described, this discrepancy can again be largely attributed to the lack of metabolic processes accounted for in *in vitro* assays. All of our mixtures contained MET, which is a prototypical "proestrogen", and thus the mixture potency was underpredicted based on the *in vitro* models. Nevertheless, BPS was more potent *in vivo* than predicted from the *in vitro* data

and the estrogenic response to the mixture of BPS + MET underscores the principle concern with relying solely on metabolically incompetent *in vitro* assays for regulatory decisions—the risk of underestimating the potency of environmental chemicals.

Our study provides cases where the use of *in vitro* assay data does not allow accurate identification of the *in vivo* activity of some compounds. This finding is consistent with other studies demonstrating a moderate to low degree of predictive capacity for estimating *in vivo* estrogenic potency of environmental chemicals from *in vitro* data and the potential for false positives and false negatives (Shen *et al.*, 2013; Zacharewski *et al.*, 1998). Shen *et al.* (2013) developed the publically available Estrogen Activity Database (EADB), which contains data for over 8000 chemicals that have been tested in various estrogen responsive *in vitro* or *in vivo* assays. In that study, the authors found a high level of agreement (>73% concordance) between *in vitro* assays; however, the agreement between *in vitro* and uterotrophic assay data was much lower (45.9, 51.3, and 56.2% concordance for reporter gene, receptor binding, and cell proliferation, respectively) with a large proportion of chemicals that were active *in vitro* eliciting no activity *in vivo*. Similarly, Zacharewski *et al.* (1998) reported that several phthalate esters were active in ER binding and transcriptional activation assays, but all were negative in the uterotrophic assay.

The studies that have reported relatively high concordance between *in vitro* ER assays and *in vivo* estrogenic responses utilized subcutaneous dosing in many of the comparisons (Akahori *et al.*, 2008; Sonneveld *et al.*, 2006; Yamasaki *et al.*, 2004; Yamasaki *et al.*, 2002). For example, Akahori *et al.* (2008) reported $R^2=0.67$ and Sonneveld *et al.* (2006) reported $R^2=0.70$ for correlations between *in vitro* estrogenic activity and *in vivo* estrogenicity. Subcutaneous dosing bypasses the physiological processes of oral absorption and first-pass hepatic metabolism. Results of *in vitro* assays are more likely to correlate well with *in vivo* uterotrophic activity following subcutaneous dosing due to bypass of gut and/or liver metabolism captured by oral dosing. This discrepancy is important because the most relevant route of exposure for xenoestrogens, particularly in humans, occurs via oral ingestion. If the goal is ultimately to determine the risk of xenoestrogen exposure to human health, *in vivo* assays should be performed with oral exposures and/or *in vitro* screening approaches should be adjusted to account for the toxicokinetics that occur following oral administration.

One approach for improving the *in vivo* predictive capacity of *in vitro* assays is incorporating metabolic competency into the *in vitro* screening methodology. For example, Punt *et al.* (2013) improved potency estimates from an *in vitro* (ERTA) screening assay to uterotrophic results by correcting for hepatic clearance of the compounds using rat liver microsome incubations of the parent compounds. However, Jacobs *et al.* (2008) discussed several challenges associated with this approach, including non-specific protein binding of the compound, target cell toxicity, and the necessity for the metabolites to be transported into the target cell. Incorporation of phase I and II enzymes into test cell lines through genetic engineering is also currently technically challenging due to decreased cell viability when too many P450s are expressed in the same cells (Coecke *et al.*, 2006). In a follow-up paper, Jacobs *et al.* (2013) described the general lack of progress in developing and standardizing methods to incorporate metabolic and toxicokinetic aspects into *in vitro* assays. The lack of metabolic competency in *in vitro* assays remains one of the most highly cited limitations in predictive toxicity methodology. Recently, several US government agencies have joined together to fund research aimed at

addressing these limitations (www.transformtoxtesting.com). The use of *in vitro* metabolizing systems to better predict *in vivo* action is a strong concept that has the potential to drastically improve the utility of *in vitro* assays for this purpose.

Moving forward, for chemicals that display questionable or equivocal estrogenicity in the current EDSP Tier 1 assays, the short-term oral *in vivo* uterotrophic assay would be useful for verifying chemical activity as an intermediate step between Tier 1 Screening and Tier 2 Testing (Gray *et al.*, 2002; Juberg *et al.*, 2014). These data would successfully identify *in vitro* false positives and chemicals that undergo significant biotransformation to more potent metabolites, or biotransformation to inactive metabolites, providing a measure of oral potency that would be useful in prioritizing chemicals and dose setting for more resource intensive Tier 2 multi-generational studies. In addition to Tier 1 screening, chemical occurrence and exposure data could be utilized to further prioritize chemicals for testing with priority given to chemicals that display high hazard as well as a potential for substantial exposure to humans or ecosystems.

The present study underscores the importance of considering the current challenges associated with using *in vitro* assay data to predict *in vivo* responses for estrogenic compounds. *In vitro* assays have unequivocal value as a screening tool to identify and prioritize compounds that have the potential to be estrogenic *in vivo* and are highly informative in determining molecular initiating events of toxic action. However, using *in vitro* screening assays without accounting for metabolism presents the risk of falsely identifying a compound as either positive or negative for activity and prevents accurate prediction/characterization of whole animal chemical potency. Further, this study shows for the first time that BPS is a more potent oral estrogenic chemical *in vivo* than is BPA and that BPC is extremely potent both *in vivo* and *in vitro*. In addition, we have quantified the degree of uncertainty in IVIVE for 7 chemicals and found that the *in vitro* potency underestimated the oral *in vivo* potency of two of them by ~65- to 305-fold. More data are needed to determine how frequently this occurs.

SUPPLEMENTARY DATA

Supplementary data are available online at <http://toxsci.oxfordjournals.org/>.

ACKNOWLEDGMENTS

The authors would like to acknowledge Paul Foster (NTP/NIEHS) and Vicki Sutherland (NTP/NIEHS) for providing some chemicals and conceptual assistance. Nicola Evans, Mary Cardon, Hunter Sampson, and Jonathan Taylor provided technical laboratory assistance.

FUNDING

Funding for this work was provided by an interagency agreement between the USEPA and the National Toxicology Program at the National Institute of Environmental Health Sciences (IA: RW-75-9228501-1) and the USEPA Safe and Sustainable Water Resources and Chemical Safety for Sustainability Research Programs.

REFERENCES

- Adams, N. R. (1995). Detection of the effects of phytoestrogens on sheep and cattle. *J. Anim. Sci.* **73**, 1509–1515.
- Ahmad, R., Verma, Y., Gautam, A., and Kumar, S. (2015). Assessment of estrogenic potential of di-n-butyl phthalate and butyl benzyl phthalate in vivo. *Toxicol. Ind. Health* **31**, 1296–1303.
- Akahori, Y., Nakai, M., Yamasaki, K., Takatsuki, M., Shimohigashi, Y., and Ohtaki, M. (2008). Relationship between the results of in vitro receptor binding assay to human estrogen receptor alpha and in vivo uterotrophic assay: Comparative study with 65 selected chemicals. *Toxicol. In Vitro* **22**, 225–231.
- Bermudez, D. S., Gray, L. E. Jr., and Wilson, V. S. (2010). Modeling the interaction of binary and ternary mixtures of estradiol with bisphenol A and bisphenol AF in an in vitro estrogen-mediated transcriptional activation assay (T47D-KBluc). *Toxicol. Sci.* **116**, 477–487.
- Bermudez, D. S., Gray, L. E., and Wilson, V. S. (2012). Modelling defined mixtures of environmental oestrogens found in domestic animal and sewage treatment effluents using an in vitro oestrogen-mediated transcriptional activation assay (T47D-KBluc). *Int. J. Androl.* **35**, 397–406.
- Bhatia, S. K., and Yetter, A. B. (2008). Correlation of visual in vitro cytotoxicity ratings of biomaterials with quantitative in vitro cell viability measurements. *Cell Biol. Toxicol.* **24**, 315–319.
- Blair, R. M., Fang, H., Branham, W. S., Hass, B. S., Dial, S. L., Moland, C. L., Tong, W., Shi, L. M., Perkins, R., and Sheehan, D. M. (2000). The estrogen receptor relative binding affinities of 188 natural and xenochemicals: Structural diversity of ligands. *Toxicol. Sci.* **54**, 138–153.
- Bonavera, J. J., Dube, M. G., Kalra, P. S., and Kalra, S. P. (1994). Anorectic effects of estrogen may be mediated by decreased neuropeptide-Y release in the hypothalamic paraventricular nucleus. *Endocrinology* **134**, 2367–2370.
- Brady, A. M., Moffat, G. J., Hall, M. G., Martens, F. K., Martens, M. A., and Nair, R. (2000). An assessment of in vivo estrogenic activity of butyl benzyl phthalate and its principal mammalian metabolites. *Toxic Subst. Mech.* **19**, 1–24.
- Bulger, W. H., Muccitelli, R. M., and Kupfer, D. (1978). Studies on the in vivo and in vitro estrogenic activities of methoxychlor and its metabolites. Role of hepatic mono-oxygenase in methoxychlor activation. *Biochem. Pharmacol.* **27**, 2417–2423.
- Coecke, S., Ahr, H., Blaauboer, B. J., Bremer, S., Casati, S., Castell, J., Combes, R., Corvi, R., Crespi, C. L., Cunningham, M. L., et al. (2006). Metabolism: A bottleneck in in vitro toxicological test development: The report and recommendations of ECVAM workshop 54. *ATLA, Altern. Lab. Anim.* **34**, 49–84.
- Cox, R. I., and Braden, A. W. (1974). The metabolism and physiological effects of phyto-estrogens in livestock. *Proc. Aust. Soc. Anim. Prod.* **10**, 122–129.
- Cummings, A. M. (1997). Methoxychlor as a model for environmental estrogens. *Crit. Rev. Toxicol.* **27**, 367–379.
- Delfosse, V., Grimaldi, M., Pons, J. L., Boulahtouf, A., le Maire, A., Cavailles, V., Labesse, G., Bourguet, W., and Balaguer, P. (2012). Structural and mechanistic insights into bisphenols action provide guidelines for risk assessment and discovery of bisphenol A substitutes. *Proc. Natl. Acad. Sci. U. S. A.* **109**, 14930–14935.
- Draganov, D. I., Markham, D. A., Beyers, D., Waechter, J. M., Jr., Dimond, S. S., Budinsky, R. A., Shiotsuka, R. N., Snyder, S. A., Ehman, K. D., and Hentges, S. G. (2015). Extensive metabolism and route-dependent pharmacokinetics of bisphenol A (BPA) in neonatal mice following oral or subcutaneous administration. *Toxicology* **333**, 168–178.
- Elshy, R., Ashby, J., Sumpter, J. P., Brooks, A. N., Pennie, W. D., Maggs, J. L., Lefevre, P. A., Odum, J., Beresford, N. A., Paton, D., et al. (2000). Obstacles to the prediction of estrogenicity from chemical structure: Assay-mediated metabolic transformation and the apparent promiscuous nature of the estrogen receptor. *Biochem. Pharmacol.* **60**, 1519–1530.
- Fan, Z., Hu, J., An, W., and Yang, M. (2013). Detection and occurrence of chlorinated byproducts of bisphenol a, nonylphenol, and estrogens in drinking water of china: comparison to the parent compounds. *Environ. Sci. Technol.* **47**, 10841–10850.
- Foster, P. M., and Gray, L. E. (2013). Toxic responses of the reproductive system. In *Casarett & Doull's Toxicology: The Basic Science of Poisons* (C. D. Klaassen, Ed.), 8th ed., pp. 861–930. McGraw Hill, New York, NY.
- Frasor, J., Barnett, D. H., Danes, J. M., Hess, R., Parlow, A. F., and Katzenellenbogen, B. S. (2003). Response-specific and ligand dose-dependent modulation of estrogen receptor (ER) alpha activity by ERbeta in the uterus. *Endocrinology* **144**, 3159–3166.
- Gray, L. E., Ostby, J., Wilson, V., Lambright, C., Bobseine, K., Hartig, P., Hotchkiss, A., Wolf, C., Furr, J., Price, M., et al. (2002). Xenoendocrine disrupters-tiered screening and testing filling key data gaps. *Toxicology* **181**, 371–382.
- Hashimoto, Y., Moriguchi, Y., Oshima, H., Kawaguchi, M., Miyazaki, K., and Nakamura, M. (2001). Measurement of estrogenic activity of chemicals for the development of new dental polymers. *Toxicol. In Vitro* **15**, 421–425.
- Heller, C. G. (1940). Metabolism of the estrogens: The effect of liver and uterus upon estrone, estradiol, and estriol 1, 2, 3. *Endocrinology* **26**, 619–630.
- Hu, Y., and Kupfer, D. (2002). Metabolism of the endocrine disruptor pesticide-methoxychlor by human P450s: Pathways involving a novel catechol metabolite. *Drug Metab. Dispos.* **30**, 1035–1042.
- Jacobs, M. N., Janssens, W., Bernauer, U., Brandon, E., Coecke, S., Combes, R., Edwards, P., Freidig, A., Freyberger, A., Kolanczyk, R., et al. (2008). The use of metabolising systems for in vivo testing of endocrine disruptors. *Curr. Drug Metab.* **9**, 796–826.
- Jacobs, M. N., Laws, S. C., Willet, K., Schmieder, P. K., Odum, J., and Bovee, T. F. (2013). In vitro metabolism and bioavailability tests for endocrine active substances: What is needed next for regulatory purposes? *Altex* **30**, 331–351.
- Jobling, S., Reynolds, T., White, R., Parker, M. G., and Sumpter, J. P. (1995). A variety of environmentally persistent chemicals, including some phthalate plasticizers, are weakly estrogenic. *Environ. Health Perspect.* **103**, 582–587.
- Juberg, D. R., Borghoff, S. J., Becker, R. A., Casey, W., Hartung, T., Holsapple, M. P., Marty, M. S., Mihaich, E. M., van der Kraak, G., Wade, M. G., et al. (2014). t4 Workshop Report: Lessons learned, challenges, and opportunities: The U.S. Endocrine Disruptor Screening Program. *Altex* **31**, 63–78.
- Judson, R., Richard, A., Dix, D. J., Houck, K., Martin, M., Kavlock, R., Dellarco, V., Henry, T., Holderman, T., Sayre, P., et al. (2009). The toxicity data landscape for environmental chemicals. *Environ. Health Perspect.* **117**, 685–695.
- Kang, J. S., Choi, J. S., Kim, W. K., Lee, Y. J., and Park, J. W. (2014). Estrogenic potency of bisphenol S, polyethersulfone and their metabolites generated by the rat liver S9 fractions on a MVLN cell using a luciferase reporter gene assay. *Reprod. Biol. Endocrinol.* **12**, 1–8.

- Kortenkamp, A. (2007). Ten years of mixing cocktails: a review of combination effects of endocrine-disrupting chemicals. *Environ. Health Perspect.* **115 Suppl 1**, 98–105.
- Kuch, H. M., and Ballschmiter, K. (2001). Determination of endocrine-disrupting phenolic compounds and estrogens in surface and drinking water by HRGC-(NCI)-MS in the picogram per liter range. *Environ. Sci. Technol.* **35**, 3201–3206.
- Kurebayashi, H., Harada, R., Stewart, R. K., Numata, H., and Ohno, Y. (2002). Disposition of a low dose of bisphenol A in male and female Cynomolgus monkeys. *Toxicol. Sci.* **68**, 32–42.
- Laws, S. C., Yavanxay, S., Cooper, R. L., and Eldridge, J. C. (2006). Nature of the binding interaction for 50 structurally diverse chemicals with rat estrogen receptors. *Toxicol. Sci.* **94**, 46–56.
- Le Fol, V., Ait-Aissa, S., Cabaton, N., Dolo, L., Grimaldi, M., Balaguer, P., Perdu, E., Debrauwer, L., Brion, F., and Zalko, D. (2015). Cell-specific biotransformation of benzophenone-2 and bisphenol-s in zebrafish and human in vitro models used for toxicity and estrogenicity screening. *Environ. Sci. Technol.* **49**, 3860–3868.
- Leet, J. K., Gall, H. E., and Sepulveda, M. S. (2011). A review of studies on androgen and estrogen exposure in fish early life stages: Effects on gene and hormonal control of sexual differentiation. *J. Appl. Toxicol.* **31**, 379–398.
- Liao, C., and Kannan, K. (2013). Concentrations and profiles of bisphenol A and other bisphenol analogues in foodstuffs from the United States and their implications for human exposure. *J. Agric. Food Chem.* **61**, 4655–4662.
- Liao, C., Liu, F., Alomirah, H., Loi, V. D., Mohd, M. A., Moon, H. B., Nakata, H., and Kannan, K. (2012a). Bisphenol S in Urine from the United States and Seven Asian Countries: Occurrence and Human Exposures. *Environ. Sci. Technol.* **46**, 6860–6866.
- Liao, C., Liu, F., Guo, Y., Moon, H. B., Nakata, H., Wu, Q., and Kannan, K. (2012b). Occurrence of eight bisphenol analogues in indoor dust from the United States and several Asian countries: implications for human exposure. *Environ. Sci. Technol.* **46**, 9138–9145.
- Liao, C., Liu, F., and Kannan, K. (2012c). Bisphenol s, a new bisphenol analogue, in paper products and currency bills and its association with bisphenol a residues. *Environ. Sci. Technol.* **46**, 6515–6522.
- Matsushima, A., Liu, X., Okada, H., Shimohigashi, M., and Shimohigashi, Y. (2010). Bisphenol AF is a full agonist for the estrogen receptor ERalpha but a highly specific antagonist for ERbeta. *Environ. Health Perspect.* **118**, 1267–1272.
- Molina-Molina, J. M., Amaya, E., Grimaldi, M., Sáenz, J. M., Real, M., Fernández, M. F., Balaguer, P., and Olea, N. (2013). In vitro study on the agonistic and antagonistic activities of bisphenol-S and other bisphenol-A congeners and derivatives via nuclear receptors. *Toxicol. Appl. Pharmacol.* **272**, 127–136.
- NAS. (2007). *Toxicity Testing in the 21st Century: A Vision and a Strategy*. The National Academy Press, Washington, DC.
- NTP. (2008). Chemical information profile for bisphenol AF [CAS No. 1478-61-1]. Supporting Nomination for Toxicological Evaluation by the National Toxicology Program. National Institutes of Health, Bethesda, MD.
- Perez, P., Pulgar, R., Olea-Serrano, F., Villalobos, M., Rivas, A., Metzler, M., Pedraza, V., and Olea, N. (1998). The estrogenicity of bisphenol A-related diphenylalkanes with various substituents at the central carbon and the hydroxy groups. *Environ. Health Perspect.* **106**, 167–174.
- Picard, K., Lhuguenot, J. C., Lavier-Canivenc, M. C., and Chagnon, M. C. (2001). Estrogenic activity and metabolism of N-butyl benzyl phthalate in vitro: Identification of the active molecule(s). *Toxicol. Appl. Pharmacol.* **172**, 108–118.
- Plowchalk, D. R., and Teeguarden, J. G. (2002). Development of a physiologically based pharmacokinetic model for estradiol in rats and humans: A biologically motivated quantitative framework for evaluating responses to estradiol and other endocrine-active compounds. *Toxicol. Sci.* **69**, 60–78.
- Pottenger, L. H., Domoradzki, J. Y., Markham, D. A., Hansen, S. C., Cagen, S. Z., and Waechter, J. J. M., (2000). The relative bioavailability and metabolism of bisphenol A in rats is dependent upon the route of administration. *Toxicol. Sci.* **54**, 3–18.
- Powers, M. S., Schenkel, L., Darley, P. E., Good, W. R., Balestra, J. C., and Place, V. A. (1985). Pharmacokinetics and pharmacodynamics of transdermal dosage forms of 17B-estradiol: Comparison with conventional oral estrogens used for hormone replacement. *Am. J. Obstet. Gynecol.* **152**, 1099–1106.
- Punt, A., Brand, W., Murk, A. J., van Wezel, A. P., Schriks, M., and Heringa, M. B. (2013). Effect of combining in vitro estrogenicity data with kinetic characteristics of estrogenic compounds on the in vivo predictive value. *Toxicol. In Vitro* **27**, 44–51.
- Richardson, S. D., and Ternes, T. A. (2005). Water analysis: Emerging contaminants and current issues. *Anal. Chem.* **77**, 3807–3838.
- Rider, C. V., and LeBlanc, G. A. (2005). An integrated addition and interaction model for assessing toxicity of chemical mixtures. *Toxicol. Sci.* **87**, 520–528.
- Rochester, J. R., and Bolden, A. L. (2015). Bisphenol S and F: A systematic review and comparison of the hormonal activity of bisphenol A substitutes. *Environ. Health Perspect.* **123**, 643–650.
- Rosenmai, A. K., Dybdahl, M., Pedersen, M., Alice van Vugt-Lussenburg, B. M., Wedeby, E. B., Taxvig, C., and Vinggaard, A. M. (2014). Are structural analogues to bisphenol a safe alternatives? *Toxicol. Sci.* **139**, 35–47.
- Sharpe, R. M. (2010). Is it time to end concerns over the estrogenic effects of bisphenol A? *Toxicol. Sci.* **114**, 1–4.
- Shen, J., Xu, L., Fang, H., Richard, A. M., Bray, J. D., Judson, R. S., Zhou, G., Colatsky, T. J., Aungst, J. L., Teng, C., et al. (2013). EADB: An estrogenic activity database for assessing potential endocrine activity. *Toxicol. Sci.* **135**, 277–291.
- Song, S., Ruan, T., Wang, T., Liu, R., and Jiang, G. (2012). Distribution and preliminary exposure assessment of bisphenol AF (BPAF) in various environmental matrices around a manufacturing plant in China. *Environ. Sci. Technol.* **46**, 13136–13143.
- Sonneveld, E., Riteco, J. A., Jansen, H. J., Pieterse, B., Brouwer, A., Schoonen, W. G., and van der Burg, B. (2006). Comparison of in vitro and in vivo screening models for androgenic and estrogenic activities. *Toxicol. Sci.* **89**, 173–187.
- Teeguarden, J. G., Calafat, A. M., Ye, X., Doerge, D. R., Churchwell, M. I., Gunawan, R., and Graham, M. K. (2011). Twenty-four hour human urine and serum profiles of bisphenol a during high-dietary exposure. *Toxicol. Sci.* **123**, 48–57.
- Tinwell, H., and Ashby, J. (2004). Sensitivity of the immature rat uterotrophic assay to mixtures of estrogens. *Environ. Health Perspect.* **112**, 575–582.
- Tominaga, T., Negishi, T., Hirooka, H., Miyachi, A., Inoue, A., Hayasaka, I., and Yoshikawa, Y. (2006). Toxicokinetics of bisphenol A in rats, monkeys and chimpanzees by the LC-MS/MS method. *Toxicology* **226**, 208–217.
- USEPA. (2004). Methoxychlor reregistration eligibility decision (RED). EPA Publication No. EPA 738-R-04-010. USEPA, Washington, DC.

- USEPA. (2009). Drinking water contaminant candidate list 3-Final. EPA-HQ-OW-2007-1189FRL-8963-6. Federal Register 74. USEPA, Washington, DC.
- Volkel, W., Colnot, T., Csanady, G. A., Filser, J. G., and Dekant, W. (2002). Metabolism and kinetics of biphenol A in humans at low doses following oral administration. *Chem. Res. Toxicol.* **15**, 1281–1287.
- Waidyanatha, S., Mathews, J. M., Patel, P. R., Black, S. R., Snyder, R. W., and Fennell, T. R. (2015). Disposition of bisphenol AF, a bisphenol A analogue, in hepatocytes in vitro and in male and female Harlan Sprague-Dawley rats and B6C3F1/N mice following oral and intravenous administration. *Xenobiotica* **45**, 811–819.
- Wilson, V. S., Bobseine, K., and Gray, L. E. Jr. (2004). Development and characterization of a cell line that stably expresses an estrogen-responsive luciferase reporter for the detection of estrogen receptor agonist and antagonists. *Toxicol. Sci.* **81**, 69–77.
- Yamasaki, K., Noda, S., Imatanaka, N., and Yakabe, Y. (2004). Comparative study of the uterotrophic potency of 14 chemicals in a uterotrophic assay and their receptor-binding affinity. *Toxicol. Lett.* **146**, 111–120.
- Yamasaki, K., Takeyoshi, M., Sawaki, M., Imatanaka, N., Shinoda, K., and Takatsuki, M. (2003). Immature rat uterotrophic assay of 18 chemicals and Hershberger assay of 30 chemicals. *Toxicology* **183**, 93–115.
- Yamasaki, K., Takeyoshi, M., Yakabe, Y., Sawaki, M., Imatanaka, N., and Takatsuki, M. (2002). Comparison of reporter gene assay and immature rat uterotrophic assay of twenty-three chemicals. *Toxicology* **170**, 21–30.
- Yang, Y., Yin, J., Yang, Y., Zhou, N., Zhang, J., Shao, B., and Wu, Y. (2012). Determination of bisphenol AF (BPAF) in tissues, serum, urine and feces of orally dosed rats by ultra-high-pressure liquid chromatography-electrospray tandem mass spectrometry. *J. Chromatogr. B: Anal. Technol. Biomed. Life Sci.* **901**, 93–97.
- Zacharewski, T. R., Meek, M. D., Clemons, J. H., Wu, Z. F., Fielden, M. R., and Matthews, J. B. (1998). Examination of the in vitro and in vivo estrogenic activities of eight commercial phthalate esters. *Toxicol. Sci.* **46**, 282–293.