Regulatory Toxicology and Pharmacology 62 (2012) 313–328

Contents lists available at SciVerse ScienceDirect

Regulatory Toxicology and Pharmacology

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

The human relevant potency threshold: Reducing uncertainty by human calibration of cumulative risk assessments

C.J. Borgert [a,b,*], E.V. Sargent [c], G. Casella [d], D.R. Dietrich [e], L.S. McCarty [f], R.J. Golden [g]

[a] Applied Pharmacology and Toxicology, Inc., Gainesville, FL, USA
[b] C.E.H.T, University of Florida, Dept. of Physiological Sciences, Gainesville, FL, USA
[c] University of Medicine and Dentistry of New Jersey, School of Public Health, Piscataway, NJ, USA
[d] University of Florida, Gainesville, FL, USA
[e] Faculty of Biology, University of Konstanz, Germany
[f] LSMcCarty Scientific Research and Consulting, New Market, Ontario, Canada
[g] ToxLogic, LLC, Potomac, MD, USA

ARTICLE INFO

Article history:
Received 16 June 2011
Available online 28 October 2011

Keywords:
Chemical mixtures
Cumulative risk assessment
Dose addition
Independent action
Phthalate esters
Anti-androgens
Endocrine disruptor
Finasteride
Diethylstilbestrol
Human relevance

ABSTRACT

The 2008 National Research Council report "Phthalates and Cumulative Risk Assessment: Tasks Ahead," rejected the underlying premises of TEQ-like approaches – e.g., chemicals are true congeners; are metabolized and detoxified similarly; produce the same biological effects by the same mode of action; exhibit parallel dose response curves – instead asserting that cumulative risk assessment should apply dose addition (DA) to all chemicals that produce "common adverse outcomes" (CAOS). Published mixtures data and a human health risk assessment for phthalates and anti-androgens were evaluated to determine how firmly the DA–CAOS concept is supported and with what level of statistical certainty the results may be extrapolated to lower doses in humans. Underlying assumptions of the DA–CAOS concept were tested for accuracy and consistency against data for two human pharmaceuticals and its logical predictions were compared to human clinical and epidemiological experience. Those analyses revealed that DA–CAOS is scientifically untenable. Therefore, an alternative approach was developed – the Human-Relevant Potency-Threshold (HRPT) – that appears to fit the data better and avoids the contradictions inherent in the DA–CAOS concept. The proposed approach recommends application of independent action for phthalates and other chemicals with potential anti-androgenic properties at current human exposure levels.

1. Background and introduction

The USEPA (EPA)[1] has established the use of relative potency approaches, i.e., toxic equivalent (TEQ)-like approaches, for mixtures risk assessment. TEQ-like approaches assume that if the following four key pharmacological or toxicological premises are met, (1) chemicals are true congeners, (2) are metabolized and detoxified by the same biological processes, (3) produce the same spectrum of biological effects by the same mode of action, and (4) exhibit parallel dose response curves for the biological effect being modeled (Safe, 1990), then one may assume that in mixtures, those chemicals will behave according to dose addition (DA) for specific toxic effects. The DA assumption treats chemicals as if they all behave as dilutions of a single prototype chemical scaled according to their potencies relative to the prototype. Thus, risks of exposure to mixtures of such chemicals are assumed to be equivalent to the risk of exposure to the total equivalent dose of the prototype chemical. Risk assessment practices at EPA and other agencies have traditionally assumed independent action (IA) for mixtures of chemicals thought to exert effects by dissimilar modes of action (ATSDR, 2001a,2001b; USEPA, 1986, 1989, 1999, 2000).

The difference between DA and IA has important practical implications for cumulative risk assessment that can be illustrated by a simple example. Consider a mixture of three nephrotoxic chemicals, each present at one-half its threshold concentration for producing tubular acidosis: IA would predict a sub-threshold effect for the mixture (i.e., 0 + 0 + 0 = 0) whereas DA would predict measurable tubular acidosis (0.5 + 0.5 + 0.5 = 1.5) (Borgert et al., 2005). Thus, IA would predict that doses of chemicals far below the observable response range would not increase the effect of other chemicals present at concentrations near or within the

* Corresponding author at: Applied Pharmacology and Toxicology, Inc., Gainesville, FL, USA. Fax: +1 352/335 8242.
E-mail addresses: cjborgert@apt-pharmatox.com (C.J. Borgert), evsargent@gmail.com (E.V. Sargent), casella@ufl.edu (G. Casella), Daniel.Dietrich@uni-konstanz.de (D.R. Dietrich), lsmccarty@rogers.com (L.S. McCarty), rgolden124@aol.com (R.J. Golden).

[1] *Abbreviations:* CAOS, common adverse outcomes; DA, dose addition; DES, diethylstilbestrol; DHT, dihydrotestosterone; EPA, Environmental Protection Agency; IA, independent action; TEQ, toxic equivalence; TDS, testicular dysgenesis syndrome.

0273-2300 © 2011 Elsevier Inc. Open access under CC BY-NC-ND license.
doi:10.1016/j.yrtph.2011.10.012

observable response range, whereas DA would predict an increased response.

The main obstacle to applying DA broadly for diverse groups of chemicals has been the required demonstration that the underlying TEQ premises are met (Safe, 1998). The latter has prompted argumentation that the TEQ concept is too restrictive for cumulative risk assessment. Indeed, the definition of DA does not include TEQ requirements, as DA is a purely quantitative model of combined action that does not require specific chemical, toxicological or pharmacological properties and requires only that the underlying dose response relationship is quantified according to a common biological metric. However, TEQ premises (Table 1) have been included explicitly to increase the reliability of extrapolating the DA model to dose ranges, chemical ratios, and species that have not been tested empirically (Safe, 1998). Hence, requiring similarity of mode of action and biological effects serves to reduce the potential for differences in the molecular, cellular and physiological response characteristics of different mixtures components from producing non-dose-additive combined action. Consequently, inclusion criteria (Table 1) reduce the potential for pharmacokinetic differences to alter DA combined action expected at molecular, cellular and physiologic levels based on similar modes of action and toxic effects, and increase confidence in treating a mixture of chemicals as if it were the total equivalent dose of the prototype chemical alone. In contrast, chemicals with non-parallel dose response curves have different relative potencies at different doses and thus, cannot reliably be treated as simple dilutions of a prototype chemical across all doses and ratios (Fig. 1).

One prominent example of how rejecting TEQ premises may lead to misinterpretation of data and over-interpretation of risk is the report "Phthalates and Cumulative Risk Assessment: Tasks Ahead" by the National Research Council (NRC, 2008). The report asserts that based on the available data for phthalates and other chemicals that result in androgen deficiency by different modes of action (anti-androgens), cumulative risk assessment should be conducted by applying DA to all chemicals that produce "common adverse outcomes" (CAOS) rather than only to chemicals satisfying TEQ criteria. Further, the report recommends DA to the exclusion of other models of combined action, such as IA, irrespective of whether chemicals exhibit parallel dose response curves for CAOS, and places no other restrictions on the application of DA to CAOS (DA–CAOS), including potency, exposure level, or mechanistic assumptions regarding the anti-androgenic effects.

Some phthalates and other chemicals with potential anti-androgenic properties have been found to produce malformations of the developing male reproductive tract in rats, an observation to which some researchers have attached the collective name "phthalate syndrome" (Foster, 2005). The mode of action for these

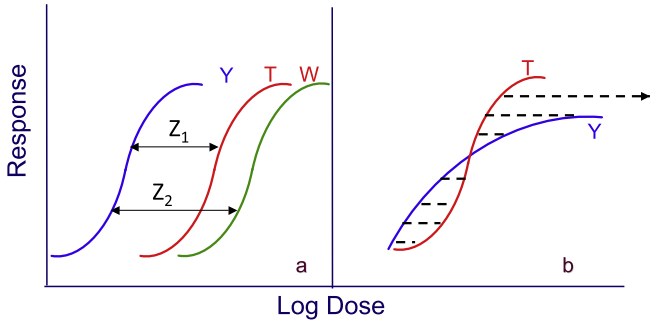

Fig. 1. Extrapolating DA requires parallel dose–response curves. In panel a, dose response curves for chemicals Y, W and T are parallel and there is a constant proportionality between the curves, illustrated by the arrows, reflecting the fact that the potency differences are constant at every response level. (Y–T; Y–W) thus give relative potency factors Z_1 and Z_2 for those chemicals, which enable prediction of the dose-additive response for untested dose combinations by transforming any dose of T or W to its equivalent dose of Y. DA Response [Y,T,W] = Response [Y + (T-Z_1) + (W-Z_2)]. These relationships hold for any number of chemicals with parallel dose response curves. In panel b, dose response curves for chemicals Y and T are not parallel, resulting in a different proportionality between the curves at every response level (dashed lines), reflecting a different relative potency Y-T at every response level. Thus, there is no uniform function that will accurately transform T into an equivalent dose of Y, and so predicting DA Response [Y,T] becomes highly uncertain for doses not tested empirically. Uncertainties expand with increasing numbers of chemicals with non-parallel dose response character.

effects is believed to involve an impairment of fetal Leydig and Sertoli cell function, leading to androgen deficiency in the developing reproductive tract of male rats (David, 2006). Similar malformations have been reported in other species, however species differences in response to phthalates are apparent (Gaido et al., 2007; Johnson et al., 2011; Lambrot et al., 2009; Hallmark et al., 2007). Based on certain similarities between these malformations in rodents and a so-called testicular dysgenesis syndrome (TDS) in humans (Skakkebaek et al., 2001), some have speculated that TDS may be caused by exposure to anti-androgenic chemicals (Skakkebaek et al., 2001; Foster, 2005).

Mixtures studies (Table 2) conducted by two research groups, one based at universities in Europe (Christiansen et al., 2008, 2009; Hass et al., 2007; Metzdorff et al., 2007) and the other at EPA (Hotchkiss et al., 2004; Howdeshell et al., 2007, 2008b; Rider et al., 2008) have reported DA combined effects on the developing male reproductive tract of rats for anti-androgens that differ in molecular structure as well as in their mechanism of action and pattern of anti-androgenic effects. Based on those findings, it has been reasoned that limiting the application of DA only to chemicals that fit TEQ criteria may be insufficiently inclusive and thus,

Table 1
Chemical similarity criteria for applying DA in cumulative (mixtures) risk assessments.

	ATSDR mixture risk assessment[a]	ILSI cumulative risk criteria[b]	EPA mixture risk assessment[c]	EPA TEQ approach[d]	DA–CAOS (NRC)[e]
Organ system (CAOS)					X
Target organ	X	X	X	X	
Molecular target	X	X	X	X	
Cellular target			X	X	
Toxic intermediates		X		X	
Pharmacokinetics				X	
Detoxification pathways				X	
Parallel dose–response curves			X	X	
Chemical structure				X	

[a] USDHHS (2001).
[b] Mileson et al. (1998) and USEPA (1999).
[c] USEPA (2000).
[d] Safe (1998).
[e] NRC (2008).

Table 2
Mixture studies of anti-androgens in rats.

Study	Chemicals	Mixture design	Model type consistent with mixture data
Hass et al. (2007)[a]	Vinclozolin Flutamide Procymidone	1 ratio 5 doses	DA fit most data; did not test IA
Metzdorff et al. (2007)[a]	Vinclozolin Flutamide Procymidone	1 ratio 5 doses	DA fit the data; did not test IA
Christiansen et al. (2008)[a]	Vinclozolin Flutamide Procymidone	1 ratio 2 doses	Did not formally test a model; concluded that results exceed IA
Christiansen et al. (2009)[a]	Diethylhexyl phthalate Vinclozolin Prochloraz Finasteride	1 ratio 3 doses	DA and IA equally fit most data; synergism observed for external malformations
Hotchkiss et al. (2004)[b]	Benzylbutyl phthalate Linuron	1 ratio 1 dose	Did not formally test a model
Howdeshell et al. (2007)[b]	Di(*n*)butyl phthalate Diethylhexyl phthalate	1 ratio 1 dose	DA and IA fit some endpoints; DA fit better than IA for most endpoints
Howdeshell et al. (2008b)[b]	Butyl benzyl phthalate Diethylhexyl phthalate Di(*n*)butyl phthalate Diisobutyl phthalate Dipentyl phthalate	1 ratio 7 doses	DA fit; Did not test IA
Rider et al. (2008)[b]	Vinclozolin Procymidone Prochloraz Linuron Butyl benzyl phthalate Diethylhexyl phthalate Di(*n*)butyl phthalate	1 ratio 4 doses	DA-based models fit data; IA did not fit data
Benson (2009)[b]	Butyl benzyl phthalate Diethylhexyl phthalate Di(*n*)butyl phthalate Diisobutyl phthalate Diisononyl phthalate Dipentyl phthalate	Risk Assessment based on common mode of action and adverse outcome	Assumed Hazard Index Approach, consistent with DA Estimated mean, upper 95% CI and maximum human exposures compared to RfDs 100-fold lower than observed mixture effects in rats (1000-fold lower for diisononyl phthalate) Concluded that exposures are within acceptable risk levels
Kortenkamp and Faust (2010)[a]	Butyl benzyl phthalate Diethylhexyl phthalate Di(*n*)butyl phthalate	Risk	Assumed Hazard Index Approach, consistent with DA
	Diisobutyl phthalate Diisononyl phthalate Vinclozolin Prochloraz Procymidone Linuron Fenitrothion p,p′-DDE Brominated diphenyl ether 99 Bisphenol A Butyl paraben Propyl paraben	Assessment based on DA–CAOS concept	Estimated mean and upper 95% CI of human exposures compared to RfDs 200–500-fold lower than observed mixture effects in rats Concluded that 95% UCI exceeds acceptable risk levels

[a] Studies conducted by the researchers at European universities.
[b] Studies conducted by the researchers at EPA.

produce insufficiently conservative human health risk assessments that fail to protect from TDS (NRC, 2008). Subsequently, two human health risk assessments have been conducted using NRC's assumptions, both of which employed a hazard index calculation to estimate risks (Table 2). One considered exposure to a mixture of six phthalate esters (Benson, 2009), while the other utilized the full DA–CAOS concept, considering exposure to a mixture of 15 potential anti-androgens spanning a range of chemical structures and modes of action including phthalate esters, pesticides, and industrial chemicals (Kortenkamp and Faust, 2010).

2. Analysis of mixtures studies, risk assessment of anti-androgens, and predictions of DA–CAOS

The DA–CAOS concept obviously contravenes well-established pharmacological principles for defining relative potency of adverse outcomes. In view of this, an in-depth analysis of the empirical and theoretical foundations of the DA–CAOS concept and the respective practical consequences of its use was carried out. As well, potency assumptions made in risk assessments based on the DA–CAOS concept were evaluated. The following analyses were conducted:

(i) An evaluation of the extent to which the study designs employed to test mixtures (Table 2) can support the extrapolation of DA versus IA to lower, untested doses and chemical ratios.
(ii) A statistical analysis of variability within the published data for one endpoint evaluated in the mixtures studies as an example to illustrate the uncertainty that might evolve from extrapolation of the DA–CAOS assumption to lower, untested doses and chemical ratios.
(iii) A consideration of whether the DA–CAOS concept can be reconciled with the pharmacological basis underlying relative potency approaches.
(iv) an evaluation of the logical extensions of the DA–CAOS concept and resulting risk assessment for consistency with human clinical and epidemiological observations, and;

(v) A comparison of humans versus rat sensitivity to chemicals with anti-androgenic properties to test the premise and key default assumption of the risk assessment based on DA–CAOS (Kortenkamp and Faust, 2010), that the developing male reproductive tract of humans responds to chemicals at doses two orders of magnitude lower than those required to affect rats.

As shown below, these five analyses demonstrated that the DA–CAOS concept and the risk assessments conducted per its premises are scientifically untenable. Consequently, an alternative approach – the Human-Relevant Potency-Threshold (HRPT) approach – was developed to better fit the data and to avoid the contradictions that arise in using the DA–CAOS concept.

2.1. Study designs

Five criteria for evaluating interaction studies used in risk assessment have been defined (Borgert et al., 2001) and were used to assess the mixtures studies listed in Table 2. Criterion 1 addresses the fact that the better defined the dose response parameters for the individual chemicals, the more reliably distinctions can be drawn between different combined effect models. Criterion 2 requires that non-interaction model(s) be explicitly defined since inferences of greater-than-additive (synergy) or less-than-additive (antagonism) are typically made on the basis of statistically significant departures from a defined non-interaction model.[2] Usually, DA and IA are the competing models of non-interaction tested. Criterion 3 requires testing an adequate number of combinations across a sufficient dose range to meet the goals of the study. The latter is important as combination effects can vary with the concentrations and with the ratios of mixture constituents. Criterion 4 requires formal statistical tests to distinguish the combined response from the response predicted by the non-interaction model. Prerequisite for a robust statistical analysis is to account for biological variation and experimental error. Finally, Criterion 5 requires the evaluation of interactions at a relevant level(s) of biological organization. Combination effects measured at molecular, cellular and higher-order physiological endpoints may be necessary to gain an unambiguous understanding of the biological response to a mixture.

2.1.1. Study designs: Dose ranges and ratios

Conducting experiments in whole animals or human subjects often prevents satisfying all five criteria (Price et al., 2002; Borgert et al., 2001) due to the difficulty of obtaining sufficient dose–response information. Those limitations affect the studies under consideration here (Table 2), as acknowledged by some of the authors (e.g., Hass et al., 2007). Nonetheless, greater conformity with the five criteria provides for more unequivocal data interpretation across the concentration ranges tested. In contrast, studies fulfilling fewer criteria or with less stringency must be interpreted with more caution, i.e., limiting the interpretation to the specific doses and ratios tested. Uncertainties are compounded when combination effects are extrapolated to doses and ratios not tested empirically.

The interpretations supportable from a mixture study are inherently dependent on the quality of dose–response data available for the individual mixture components, which should be tested across a concentration range and with a sufficient number and spacing of doses to reveal maxima, minima, points of inflection, and regions of linearity (Borgert et al., 2001). The mixtures studies (Table 2) used for the risk assessment of anti-androgens (Kortenkamp and Faust, 2010) varied widely in their characterization of dose response data for the individual chemicals. In some studies, as many as seven doses were tested while in other studies, dose response data from prior experiments and different rat strains were used as surrogates. In most instances, dose response models were used to curve-fit the data, and model parameters obtained from single chemical experiments were used to predict mixture effects.

All of the studies employed single ratios of chemicals to simulate mixture effects (Table 2). This design is often referred to as a fixed ray design (Cassee et al., 1998) and has advantages over experimental designs employing only a single concentration of the mixture components because it allows local interpretations beyond one data point. A fixed ray design may be the broadest study design achievable in live animals due to limitations on the manageable size and number of dose groups (Cassee et al., 1998; Price et al., 2002), but this feature should constrain the interpretations to the ratios tested and preclude extrapolation to untested ratios. Several of the studies chose individual constituent ratios predicted to yield an equal contribution from each component across the entire range tested, however, confidence in that prediction is unjustified unless all components have parallel dose response characteristics. It is highly questionable whether the slopes of the dose response data for male reproductive tract effects of the individual chemicals are sufficiently similar to support this assumption. Although DA assumptions and calculations can be made for chemicals with non-parallel dose response curves, the reliability of those calculations diminishes rapidly as they are extrapolated beyond the ratios and concentrations tested empirically (Fig. 1) (Cassee et al., 1998; Borgert et al., 2001; Price et al., 2002).

Although the mixtures studies (Table 2) all reported testing "low doses" of the mixture components administered to dams, this term must be understood in the context of the physiological system. The doses tested appear to be within an order of magnitude of the observable response range for physiologically relevant anti-androgenic effects. For example, Rider et al. (2008) reported that the mixture doses used in their study were below the observable response range for malformations of the developing male reproductive tract in rats, i.e., procymidone at a maternal dose of 7.5 mg/kg/day. Contrary to the latter no observable effect assumption, Metzdorff et al. (2007) reported that 5 mg/kg/day procymidone produced statistically significant changes in seminal vesicle weight, and 10 mg/kg/day produced changes in testis, ventral prostate, levator ani/bulbocavernosus muscle, and bulbourethral gland weights, thus suggesting that Rider et al. (2009) were testing mixtures containing doses of procymidone well within the observable effect range for anti-androgenic action. Further, although Rider et al. (2008) reported that vinclozolin was without statistically significant effects at 3.75 mg/kg/day, Metzdorff et al. (2007) reported that a slightly higher dose of 10 mg/kg/day vinclozolin produced statistically significant changes in epididymal, ventral prostate, and seminal vesicle weights. These examples demonstrate the substantial variability that exists in defining the observable response range, especially for endocrine-sensitive endpoints (Ashby, 2003). Because the comparison of no-interaction dose–response models for mixtures, i.e., DA versus IA, entirely depends on the precision of the no observable effect estimate, this latter precision must influence the confidence placed on the interpretations drawn from such data, as exemplified below.

Rider et al. (2008) also showed that individual phthalates (dibutyl phthalate and diethylhexyl phthalate) failed to increase the incidence of male reproductive tract malformations at maternal doses below 500 mg/kg/day,[3] but that in a mixture with the anti-

[2] Models that test directly for synergism have been devised (e.g., Barton et al., 1993; Laska et al., 1994) but have not gained wide acceptance.

[3] Ranges inferred from Fig. 1 of the cited paper.

androgenic pesticides vinclozolin, procymidone, prochloraz, and linuron, doses of 75 mg/kg/day of butyl benzyl phthalate, dibutyl phthalate and diethylhexyl phthalate contributed to an increased incidence of observable malformations.[4] This suggests that the combination of the three phthalates (75 mg/kg/day each, thus a total of 220 mg/kg/day) would provide for an effect greater than the no observable effect estimate demonstrated for 500 mg/kg/day of dibutyl phthalate or diethylhexyl phthalate. In a previous study by the same group, these three phthalates reduced fetal testosterone production at doses of 300 mg/kg/day individually and at doses of 60 mg/kg/day in combination with two other phthalates (Howdeshell et al., 2008b), suggesting that the three phthalates alone (butyl benzyl phthalate, dibutyl phthalate and diethylhexyl phthalate) would reduce fetal testosterone production when combined in a mixture at approximately 100 mg/kg/day each. Thus, although the doses of phthalates and other chemicals used in these mixtures were statistically below their individual no effect levels on male reproductive tract malformations in particular studies, they are nonetheless close to the dose range that produces a clear reduction in fetal testosterone individually and in mixtures. Slight differences in the experimental protocols, time of dosing, or rat strain could explain these differential responses. Regardless of the underlying reason, this comparison underscores how imprecise the distinctions might be regarding the no observable effect level for androgen sensitive tissues, and illustrates that the label "low dose" cannot be taken to mean a dose that is without a physiologically relevant anti-androgenic effect. To put the doses used in these rat studies into perspective with human-relevant exposures, the "high intake" level of vinclozolin and procymidone (9 μg/kg/day) and butyl benzyl phthalate (4 μg/kg/day), dibutyl phthalate (6 μg/kg/day), and diethylehexyl phthalate (3.6 μg/kg/day) estimated for the US population are roughly 3–4 orders of magnitude lower (Kortenkamp and Faust, 2010).

Despite limitations just described, the DA–CAOS concept was extended well below the dose range where DA was demonstrated (Kortenkamp and Faust, 2010; NRC, 2008). Fig. 2a illustrates the general study designs that have produced DA mixture effects from "no-effect" combinations of anti-androgenic chemicals with different modes of action. Fig. 2b illustrates how the DA–CAOS approach extrapolates the same studies depicted in Fig. 2a to far lower doses of the mixture components, which were not tested empirically. The conservatism introduced here goes well beyond extrapolating observed toxicity from high to low doses; that conservatism is compounded by the choice of DA over IA based on very limited study designs.

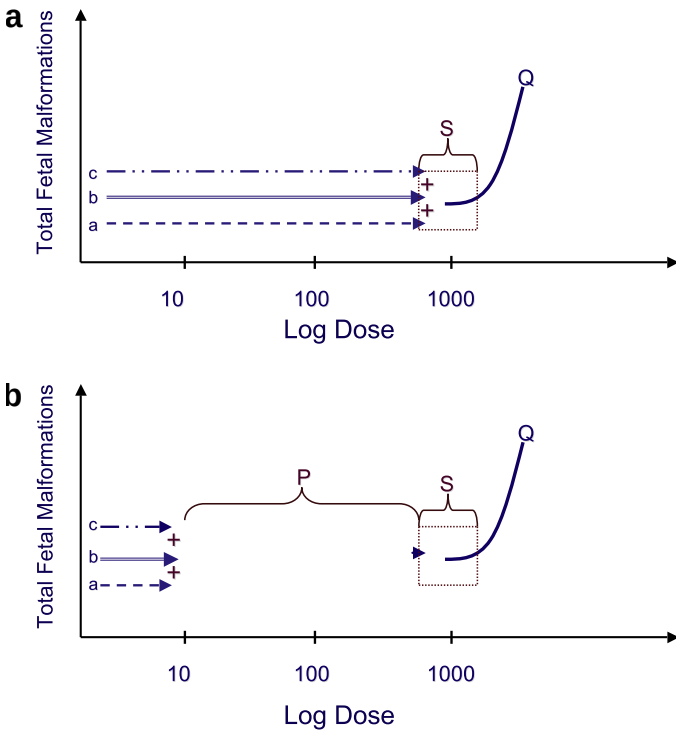

Fig. 2. Conceptual Model for DA–CAOS. Different line patterns a, b, and c represents chemicals with anti-androgenic potential that produce CAOS via different modes of action. (S) indicates the dose range where mixture experiments (Table 2) have been performed relative to the dose–response curves for observable effects (Q). Panel a: Generation of dose addition data. Available data suggest that sub-effective doses of a few such chemicals may produce a DA response when the dose of each is within S, i.e., near the observable response region Q. Panel b. Extrapolation of mixture data to DA–CAOS model. The DA–CAOS concept assumes that doses of such chemicals far below the observable response region (S) also operate by DA and may produce an observable response if sufficient numbers of chemicals are present. However, no data support this extrapolation (P).

2.1.2. Study design: Endpoints and dose response metrics

Mixtures studies on potential anti-androgenic chemicals consistently report conformity with the DA model of combined action, however, those studies report inconsistent results as to whether the data also conform to IA (Table 2). For example, one study reported that IA under-predicted combination effects and only DA–models adequately fit the data for all malformations combined (Rider et al., 2008), whereas another study (Christiansen et al., 2009) found that the data for most endpoints could be fit adequately by either model, with some greater-than-additive exceptions. The inconsistency of results obtained from the different approaches used by these two research groups could be due to any number of factors, including the animals used, the exact doses tested, the way endpoints were measured, or slight differences in the mathematical algorithms used for DA and IA.

Although the application of DA is not constrained to any particular type of effect, the use of scored endpoints, which are inherently subjective (Haschek et al., 2010), presents challenges for analyzing experimental variance not typically encountered when continuous variables are measured according to objective scales. Several of the mixtures studies listed in Table 2 assessed scored endpoints, but it is unclear how variance was assessed statistically for these endpoints, if it was addressed at all. Except for gubernacular underdevelopment (not categorized as a malformation above a certain length), all other male reproductive tract malformations were combined into a single group, further complicating the assessment of experimental variance; it is unclear if or how this was addressed in the statistical analysis (Howdeshell et al., 2007; Rider et al., 2008). Other studies categorized fetal malformations according to a four-point scale that included none observable, mild, moderate, or severe (Christiansen et al., 2008, 2009). To reduce those scores to a dichotomous variable suitable for statistical analysis used to test DA versus IA combined effect models, moderate and severe malformations were grouped together in the "malformations" category, while mild malformations were grouped together with no observable malformations in the "no-effect' category. Although this practice allows some statistical analysis of results, it introduces additional potential errors of interpretation.

IA predicts that no-effect doses of individual chemicals will also produce no effect when combined. Therefore, including mild malformations or gubernacular underdevelopment in the no-effect category for single chemical responses, as was done in some analyses (Christiansen et al., 2008, 2009) increases the chance that malformations will be observed when so-called "no-effect" doses of several chemicals are combined in the mixtures study. Because malformation severity also increases with dose, these methods of

[4] Inference from Fig. 3 of the cited paper.

scoring and grouping malformations may have ensured rejection of the IA model or biased the analysis toward synergism because a slight increase in "dose" from the combination of agents would raise certain mild malformations to the moderate/severe category. The uncertainty introduced by this procedure was not addressed, and methods for assessing its impact are lacking. The fact that small differences in experimental protocol or analysis alter the results of the mixture experiment raises concerns as to the degree of uncertainty inherent in interpreting the results for risk assessment. The risk assessment of anti-androgens chose DA over IA based on a stated preference for the more conservative model (Kortenkamp and Faust, 2010), however, the implications of inconsistent results for extrapolating to untested doses and mixture ratios appear to have been overlooked.

2.2. Statistical analysis of variability

In order to provide an objective estimate of the potential uncertainty contributed by the dose response information discussed above, data presented in one of the published mixture studies (Rider et al., 2008) was evaluated. Because the original raw data were unavailable, data points were inferred from published figures (Rider et al., 2008, Fig. 1). In order to be as consistent as possible with the published study, the dose response model used to fit the published data (Rider et al., 2008) was also used in the analysis presented here. Specifically, based on the data points inferred from the graphs published by Rider et al. (2008) and the dose response model they fit to their data, a dose–response curve was developed and then data points were generated from the theoretical curve. A number of different samples were obtained in this way. Such iterations of the dose–response experiment comprise a statistical bootstrap procedure (Efron and Tibshirani, 1993), and the results are presented in a series of isobolograms (Fig. 3). The variability in the bootstrap samples is representative of that in the original data, providing a clear picture of the range of doses consistent with any particular level of response.

Isobolograms are a simple means of graphically evaluating data on binary mixtures for conformity to DA. Doses of one of the mixture components are plotted along the abscissa and the other along the ordinate. The equation for DA describes a line connecting equally effective doses of these two chemicals on the ordinate and abscissa. All other equally effective doses, representing defined mixtures of both components, are DA if they fall on the line, less than DA if they fall above the line, and greater than DA if below the line. Since some degree of variability and experimental error are inherent to the measurement of any observation or biological endpoint, the lines of additivity representing DA combination doses in an isobologram must be enveloped by statistical

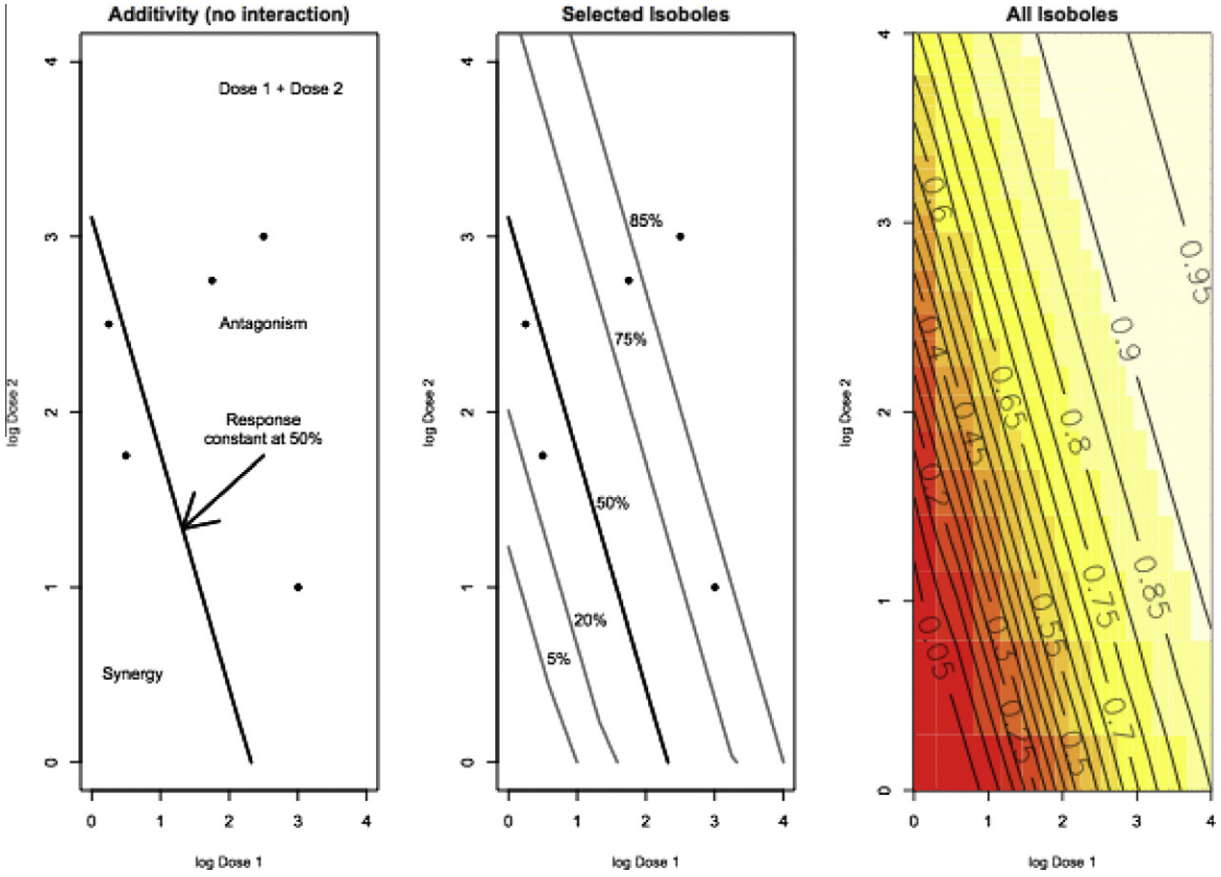

Fig. 3. Statistical analysis of uncertainty; isoboles created from Bootstrap procedure. 3–1: The left panel is an isobologram for two agents, Dose 1 and Dose 2, at 50% response. Dose addition (DA) defines a line connecting equally effective doses of Dose 1 and Dose 2 administered individually. Points above the line would demonstrate less-than-additive (antagonistic) dose combinations, below the line, greater-than-additive (synergistic). The dose is plotted on a log 2 scale, so a change from 2 to 4 on the log 2 scale would represent a doubling of the dose. The center panel shows a collection of isoboles at different response levels. Again, each line represents a constant response over the mixture of doses. The right panel shows the output from one analysis, and depicts all of the isobolograms, from 5% to 95% response. As this is the output from one sample, the isobolograms will vary from sample to sample. 3–2: This variation is illustrated for the data inferred from Fig. 1 of Rider et al. (2008). The estimated variance in equi-effective doses, and thus, the extent of uncertainty surrounding the DA prediction, is illustrated by comparing how the 95% effective dose for one mixture component changes across experiments as plotted on the abscissa. The lowest log (base-2) dose estimated to produce a 95% response is approximately 1.7 (panel e) whereas the highest dose estimated to produce the same level of response is approximately log dose 3.2 (panel c), indicating that the iso-effective dose of just one mixture component can vary nearly 2-fold even within the observed data. The parallel lines correspond to isoboles for equi-effective doses at lower response levels indicate that the variance at lower doses and response levels is proportional to the variance at higher doses and response levels, i.e., nearly twofold.

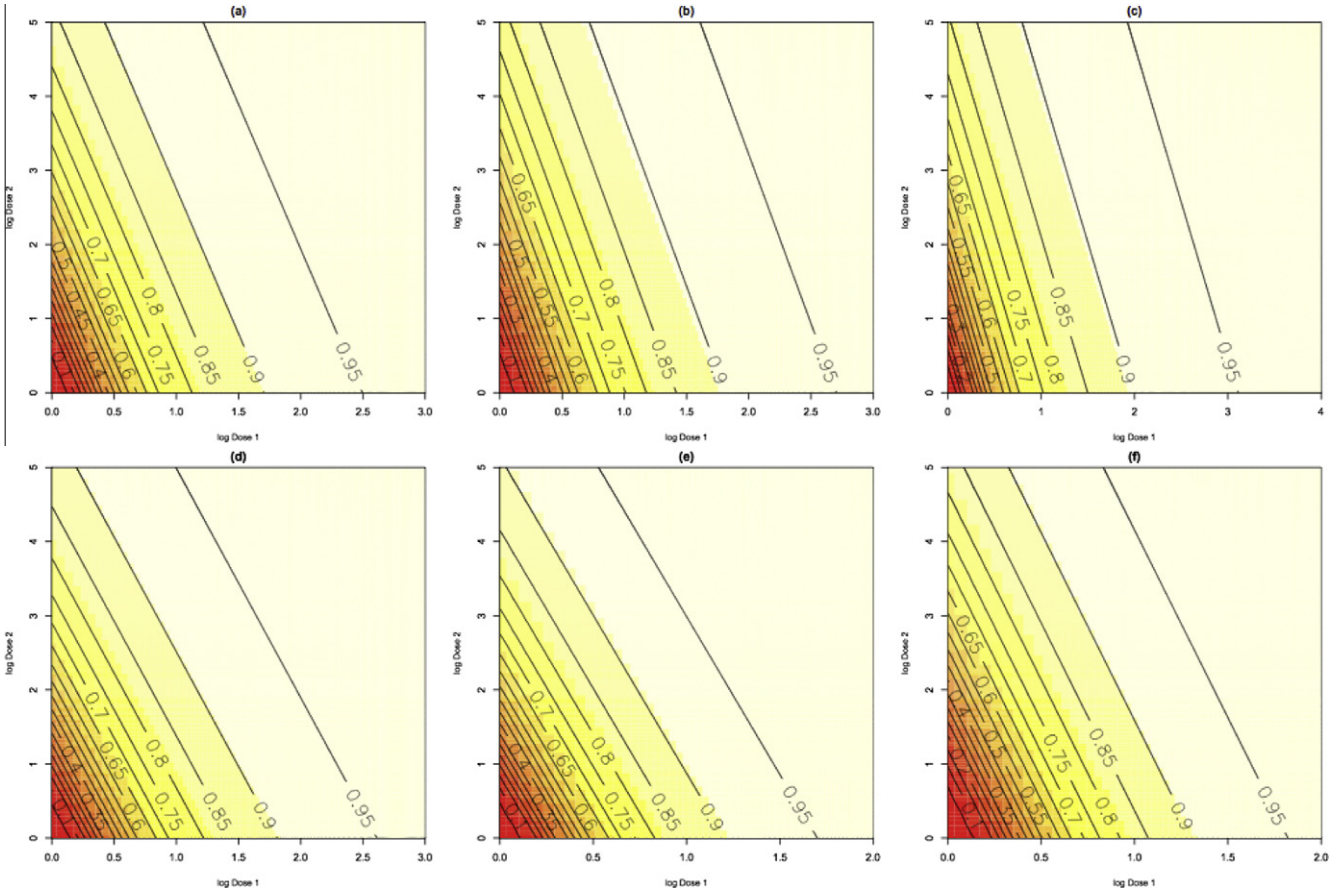

Fig. 3 (*continued*)

confidence intervals defined by this variance. Assays with higher variance and/or experimental error will produce larger confidence intervals than assays with lower variance and/or experimental error (Borgert et al., 2005).

The variability attending any DA prediction can be estimated by observing how the equi-effective dose of each mixture component varies from experiment to experiment. The isobolograms shown in Fig. 3(2), a–f indicate considerable variance within the published data, despite the fact that the bootstrap procedure used to generate the new "data" on which the isoboles in Fig. 3(2) were constructed employed the underlying dose–response model that Rider et al. (2008) concluded best fit their data. Even within these contrived, best-case experiments, nearly 2-fold variance is observed, indicating that this variability surrounds any DA conclusion, even within the range of doses tested empirically. It is thus surprising that Rider et al. (2008) unequivocally excluded IA for nearly all of the reported combination effects, since this requires an absence of overlap between IA and DA predictions after accounting for the experimental variance. It is unclear how this variance was handled in the analysis of combination effects (Rider et al., 2008). Similar to the analysis presented here, other studies employed a bootstrap procedure to assess variance, and on that basis, concluded that both IA and DA predicted most of the mixture responses (Metzdorff et al., 2007).

It is critical to appreciate that for assessing risk, the DA–CAOS theory extrapolates the DA model beyond the doses and ratios tested in mixtures studies (Kortenkamp and Faust, 2010). Since the apparent variance of the data within the range of doses tested in those studies (Table 2) raises questions as to whether a single model of combined action can be unequivocally declared the most accurate, even for all mixture ratios within that dose range, extrapolation to much lower doses and different mixture ratios would be quite tenuous. Little attention was given to the fact that the uncertainty of the predictions will expand in accordance with the variance in the data, other than to justify the choice of the model on a preference for conservatism.

The statistical analysis presented here has implications for the feasibility of selecting a 'best' model of combined action as well as for future mixtures studies aimed at supporting such a determination. Unless the observed variance were so small that no significant overlap of DA and IA model predictions could occur as the model is extrapolated across untested dose ranges, it would be futile to attempt to select a model of combined action based on a statistical analysis of mixtures data within the observed response region. Indeed, the research group from Europe found that DA and IA overlapped even within the dose regions tested (Table 2), substantiating this point.

2.3. Relative potency and pharmacokinetics: Pharmacological principles

In addition to the problems discussed above, the DA–CAOS theory contravenes fundamental pharmacological principles of receptor/enzyme affinity, intrinsic activity, and potency. Affinity is a term referring to the strength of attachment between two molecules and is applicable to interactions of a small ligand with a larger macromolecule such as an intracellular or membrane-bound receptor, an ion channel, an enzyme, or specific binding protein. For simplicity, all of these are referred to as 'receptors' with the understanding that the general principles apply to all such interactions with macromolecules. The affinity of a ligand for a specific receptor determines its residence time of association, a parameter

often quantified by the dissociation constant. Generally, higher affinity ligands have longer residence times. Intrinsic activity is the relative ability of a drug-receptor complex to produce a maximum functional response and is sometimes used interchangeably with efficacy. However, 'intrinsic activity' refers to a cellular response whereas 'efficacy' is more often used in the context of a clinical response. Assuming equivalent pharmacokinetic parameters and affinity, a drug with greater intrinsic activity would have greater efficacy. Potency is the intensity of effect produced per unit of drug, and is a function of intrinsic activity and affinity.

Because most receptor-based physiological responses can be triggered when only a small fraction of available receptors are activated by a strong agonist (a ligand with high affinity and intrinsic activity), receptor ligands with very low affinity and no intrinsic activity (weak antagonists) will fail to interfere with endogenous agonists unless their concentrations reach levels that obstruct access to the receptor by shear mass action. The same principle applies to competitive inhibitors of enzymes involved in steroidogenic pathways or to activators or blockers of ion channels in cellular membranes. Thus, because of vastly different residence times, ligands with very low affinity spend such little time in contact with a receptor that they produce no discernible interference with high-affinity ligands unless their concentrations reach a sufficient level that interference by mass action occurs. In other words, low affinity ligands cannot compete with high affinity ligands by their strength of attachment, but rather, only by shear numbers of molecules that impede access to the receptor. This is why low affinity ligands have no discernible effects at low concentrations. Even agonists with relatively similar affinity but different intrinsic activity are incapable of producing linear isoboles, commonly used to detect DA (see Section 2.2 above), because the physiological response reflects both affinity and intrinsic activity. Drugs or chemicals that differ in intrinsic activity will necessarily compete at a molecular site of action and will not produce a linear DA combined response, as demonstrated mathematically and empirically by Tallarida (2006, 2007).

In contradiction of these principles, DA–CAOS posits, for example, that any dose of a weak androgen receptor antagonist will diminish the physiological activity of dihydrotestosterone by some finite degree, irrespective of whether its affinity approaches that of a natural ligand. This demonstrably faulty premise provides the underlying basis for the DA–CAOS supposition that chemicals capable of reducing androgen levels at high doses will add to the effect of weak receptor antagonists even at very low doses. In fact, however, multiple chemicals given at doses incapable of affecting physiological processes individually, by separate mechanisms, would not be expected to produce a physiological effect in combination. A combined physiological effect would not be expected unless doses of those chemicals were on the cusp of producing overt effects individually. In other words, fundamental pharmacological principles dictate that a weak androgen receptor antagonist would not produce a combination effect with an androgen synthesis inhibitor or Leydig cell toxicant unless the doses were sufficient to reduce both the concentration of endogenous androgen and the numbers of available receptors to levels near the critical minima necessary for supporting normal physiology. The DA–CAOS theory fails to recognize that the presence of myriad weak hormone receptor agonists and antagonists in the environment would fail to be DA in combination or to achieve physiological significance, in part because their weak properties cancel one another and in part because their low affinities and intrinsic activities preclude it (Safe, 1998; Tallarida, 2006).

Furthermore, the role of pharmacokinetics and how these phenomena may change with dose and ratio of mixture constituents is often under-appreciated in discussions of the joint toxicity of chemicals. This is important because pharmacokinetic alterations underlie the majority of documented interactions between chemicals (Krishnan and Brodeur, 1994) and the dose-dependence of many toxicity mechanisms, even for single chemicals (Slikker et al., 2004). Obviously, the potential for concomitant administration to affect the absorption, distribution, metabolism and excretion of chemicals in mixtures increases with dose and the number of chemical constituents as the underlying processes that control each reach the limits of their capacity. Such influences would be of greatest consequence for highly potent compounds. Without understanding how pharmacokinetic processes change with dose and ratio of constituents, it is impossible to reliably extrapolate combination models across different dose ranges and chemical ratios. Neither the mixture studies of potential anti-androgens (Table 2) nor the NRC report (NRC, 2008) considered these issues.

2.4. Testing predictions of the DA–CAOS concept

Despite the uncertainties inherent in the published mixtures studies and the unnecessary conservatism introduced into the risk assessment by assuming DA at all chemical concentrations, it is important to consider whether the logical extensions of the DA–CAOS theory are nonetheless concordant with human clinical and epidemiological experience. To test this concordance, an attempt was made to reconcile the incidence of TDS in humans with application of the DA–CAOS concept to the full suite of chemicals alleged to have anti-androgenic potential. As well, the DES dose–response for male reproductive tract malformations was considered according to the DA–CAOS concept in light of extant information regarding other environmental exposures during the period when DES was administered to pregnant women.

2.4.1. Incidence of TDS and cumulative exposure to anti-androgens

The published risk assessment based on the DA–CAOS concept (Kortenkamp and Faust, 2010) concludes that "*...the cumulative risks from anti-androgen exposures exceed acceptable levels for people on the upper end of exposure levels*," i.e., the upper 95% confidence interval of human exposures to only 15 anti-androgenic chemicals. This conclusion implies that pregnant women and their fetuses are at an unacceptably high risk for anti-androgenic effects, and specifically, that approximately 5% of male fetuses are at risk for development of TDS. The authors also claim that 8% of all known chemicals are likely to possess anti-androgenic potential, including thousands of chemicals on the market in the European Union (Kortenkamp and Faust, 2010). Presumably, similar exposures occur in other industrialized nations. Logically extending the DA–CAOS theory would thus project that the percentage of human fetuses at risk for the development of TDS is actually much higher than 5% if exposure to all chemicals with anti-androgenic potential were included. Given the author's contention that thousands of chemicals marketed in Europe may have anti-androgenic potential, their published risk assessment on 15 anti-androgens considered only 1% or fewer of the relevant chemicals. If one further considers that in utero exposures to over-the-counter analgesics have also been linked to similar male reproductive tract disorders in rats and humans (Kristensen et al., 2011), the projected percentage of affected fetuses should approach 100%, depending on exposure levels for the various putative anti-androgens.

However, the actual incidence of TDS could not be nearly so high, as the incidence of hypospadias and cryptorchidism have been estimated at between 0.2–1% and 2–9% respectively (Toppari et al., 2010). Even without considering the DA–CAOS theory, some clinicians have questioned the etiologic role of industrial chemicals in TDS (Thorup et al., 2010), asserting that the epidemiological data do not support such a relationship. Their in-depth analysis of the incidence and biology of these male reproductive tract

abnormalities argues against the existence of a 'syndrome' of effects caused by a common etiologic agent(s), instead pointing to a complex array of clinical diagnostic, genetic, and other factors that may be individually involved in the apparent increased incidence of hypospadias, cryptorchidism, testicular cancer and other malformations (Thorup et al., 2010). For example, an analysis by Fisch et al. (2001) revealed that hypospadias was significantly associated with increasing maternal age as a consequence of more women who delay childbearing until their mid-30s. It is also worth noting that Sharpe (2003), who was initially one of the principle proponents of the TDS hypothesis as a consequence of in utero/neonatal exposure to a variety of weakly estrogenic compounds from the environment, offered the following reassessment: "What is reasonably clear is that all of the identified "environmental estrogens" possess weak or very weak intrinsic estrogenic activity when measured by conventional in vitro and in vivo assays for estrogenicity... By comparison with the potency of DES, for which there [are] both human and rodent data on incidence of male reproductive developmental disorders following in utero exposure (or neonatal exposure in rodents), it seems unlikely that any of the identified environmental compounds could induce either cryptorchidism, hypospadias or testis germ cell cancer and only a tiny possibility that such compounds could affect sperm counts/sperm production. ...Based on estrogenic potency, human exposure to the most potent environmental estrogens would need to be at least 1000-fold higher than this level for adverse effects relevant to the human male to be induced, and such levels of exposure are remote."

2.4.2. Cumulative exposure to anti-androgens and clinical threshold for DES

Because normal development of the male reproductive tract is dependent on the androgen/estrogen ratio, the DA–CAOS theory predicts that cumulative exposure to environmental estrogens, anti-androgens, and other chemicals that can produce TDS-like abnormalities will increase its incidence in a DA manner. Although great concern has been generated over current cumulative exposures to anti-androgenic chemicals, it is generally acknowledged that human exposures to a wide array of chemicals capable of affecting the developing male reproductive tract – both anti-androgenic and estrogenic – have decreased since the period during which diethylstilbestrol (DES) was administered to pregnant women and its eventual removal from the market, in part for its TDS-like effects induced in utero [see Section 2.5.1 for a detailed discussion]. Polychlorinated biphenyls (PCBs) (Andric et al., 2000a,b; Gray et al., 1999) chlorinated pesticides (Fernandez et al., 2007; Kelce et al., 1995), and 2,3,7,8-tetrachloro-p-dioxin (TCDD) (Gray et al., 1995, 1997) are but a few prominent examples of chemicals with anti-androgenic potential to which human exposures have been declining since the DES episode (Adeshina and Todd, 1991; Axmon et al., 2008; Hays and Aylward, 2001; Hovinga et al., 1992; Petreas et al., 2001; CDC, 2005; USEPA, 2006). Given that use of high-dose over-the-counter analgesics, which have also been associated with male reproductive tract malformations (Kristensen et al., 2011), was also common during the era when DES was given, it seems highly probable that exposures to chemicals capable of producing TDS-like effects were substantial during the DES episode, and most certainly higher than today.

If, as the DA–CAOS risk assessment predicts, a significant proportion of male fetuses experience anti-androgenic effects today, at least as high a proportion would likely have experienced such effects during the era of DES use in pregnancy. Assuming the DA–CAOS theory, it would follow that any dose of DES should have produced an observable increase in the incidence of TDS-like effects above the predicted observable background of chemical-induced male reproductive tract malformations. In contrast, no clear increase was observed in the incidence of TDS-like effects with the lower-dose regimens of DES (Dietrich, 2010; Golden et al., 1998).

2.5. Human versus rat sensitivity

One reason the hazard index-based DA–CAOS risk assessment (Kortenkamp and Faust, 2010) is irreconcilable with human clinical and epidemiological evidence is it employed reference doses (RfDs) that were developed without considering relevant data and physiological knowledge concerning species-specific sensitivity for chemical effects on the male reproductive tract (Cook et al., 1999; Hallmark et al., 2007; Scott et al., 2009). The derived RfDs incorporate uncertainty factors of 200–500 (Kortenkamp and Faust, 2010), consistent with an assumption that adverse effects on the developing male reproductive tract may be observed in humans at doses 200–500-fold lower than doses required to elicit such effects in rats. In the absence of relevant human data, such procedures are considered appropriate for extrapolating rodent toxicity data to humans. However, direct human data are available for many human pharmaceuticals, including drugs with anti-androgenic and/or TDS-like effects. The assumptions made in the risk assessment were thus tested against such human data. Data were evaluated by comparing human versus rat data from (a) in utero exposures to DES, a chemical that produces TDS-like malformations of rodent and human reproductive tracts by inducing androgen deficiency secondary to functional disruption of Leydig and Sertoli cells, and (b) human versus rat administration of finasteride, an anti-androgen that produces androgen deficiency by inhibiting conversion of testosterone to its active form, dihydrotestosterone.

2.5.1. Human versus rat – DES

To test the premise that chemicals affect the developing male reproductive tract of humans at doses 200–500 lower than in rats, human clinical data on gestational exposure to DES were compared with data from concordant exposures in the rat. Although DES is a potent estrogen agonist that interferes with Sertoli and Leydig cell function in rodents, some of its adverse effects on the male reproductive tract are complex, e.g., may involve both estrogen (Couse and Korach, 2004) and androgen pathways (Goyal et al., 2009; Rivas et al., 2003), and exhibit both similarities and differences compared to pure estrogens such as estradiol 17β (Adachi et al., 2004; Khan et al., 1998; Lassurguere et al., 2003; Warita et al., 2010). In addition to reducing activity of the steroidogenic acute regulatory protein (StAR) and other effects similar to those produced by some phthalates in rats (Guyot et al., 2004; Howdeshell et al., 2008a; Ikeda et al., 2008), treatment of neonatal rats with testosterone or dihydrotestosterone prevents most effects of DES on the developing male reproductive tract (Rivas et al., 2003; Goyal et al., 2009), providing evidence that androgen deficiency or alteration of the androgen-estrogen balance is involved in DES action. Loss of Leydig cell function, also proposed for phthalate esters (David, 2006; Howdeshell et al., 2007), is one of many anti-androgenic modes of action that can result in androgen deficiency and is encompassed by the DA–CAOS concept. Thus, the action of phthalate esters on fetal Leydig and Sertoli cells appears to be more similar to DES than to the other anti-androgens assessed in the mixtures studies analyzed here (Table 2).

Beginning in the 1940s, DES was widely prescribed to some 5 million pregnant women under the mistaken assumption that it prevented miscarriage. The discovery that gestational exposure to DES induced a low incidence of clear cell vaginal adenocarinoma in daughters and a low incidence of male reproductive abnormalities in sons led to its removal from the market in 1972. Its effects on sons exposed in utero, including epididymal cysts,

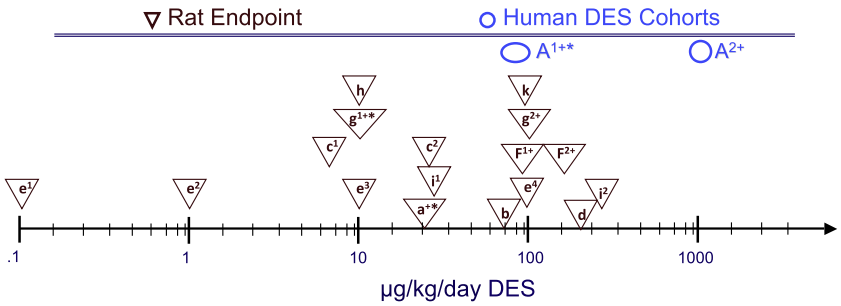

Fig. 4. DES potency comparison for male reproductive tract parameters. Human clinical data (ovals) Rat experimental data (triangles). Asterisks (∗) denote no-effect doses. Plus (+) denotes in utero administration. Pregnant women were assumed to weigh 70 kg. Where doses to rats were not reported per body weight, body weight data from Klinger et al. (1996) or Pullen (1976) were used to calculate approximate administered doses. (a) Adamsson et al. (2008). In utero exposure to Sprague–Dawley dams. +*No effect on fetal testicular T, Prog, StAR protein (steroidogenic acute regulatory protein), AR protein expression. (b) Filipiak et al. (2009). Administration on postnatal days 5–15 to Wistar pups. Reduced testes relative weight, seminiferous tubule diameter and length at puberty. (c) Goyal et al. (2001). Administration to 70–80 day old adult Sprague–Dawely males for 12 days. (1) Markedly reduced plasma testosterone levels; (2) reduced size and number of Leydig cells and plasma testosterone barely detectable. (d) Goyal et al. (2004). Administration every other day from postnatal days 2–12 to Sprague–Dawley pups: reduced penis weight, size and altered morphology; plasma testosterone levels undetectable. (e) Goyal et al. (2005). Administration every other day from pnd 2–12 to Sprague–Dawley pups; dose calculated based on 10 g rat pup average, and not averaged over days, i.e., plotted doses are overestimates. (1) Reduced weight of caudal epididymal fat pad (2) Reduced weights of caudal epididymal fat pad and seminal vesicles. (3) Reduced weights of caudal epididymal fat pad, seminal vesicles, testis, and reduced penis diameter. (4) Reduced weights of caudal epididymal fat pad, seminal vesicles, testis, and reduced penis diameter, weight and length. (f) Haavisto et al. (2001). In utero administration on embryonic days 13.5, 15.5 and 17.5 to Sprague–Dawley dams. +(1) 50% reduction in fetal plasma and testicular testosterone levels. +(2) Reduced hCG-stimulated testosterone surge. (g) Haavisto et al. (2003). In utero administration on embryonic days 13.5, 15.5 and 17.5 to Sprague–Dawley dams. +*(1) No-effects on fetal testicular and plasma testosterone. +(2) Reduced fetal testicular and plasma testosterone. (h) Mathews et al. (2009). Administration on postnatal days 1–6 to male Sprague–Dawley pups. Reduced testes weight and altered epididymal morphology, reduced androgen receptor expression and Leydig cell volume. (i) McKinnell et al., 2001. Administration to male Wistar rat pups every other day on postnatal days 2–12. (1) 38% reduction testes weight. (2) Reduced testes weight and altered epididymal morphology, reduced androgen receptor expression and 91% reduction in Leydig cell volume. (k) Mikkilä et al. (2006). Subcutaneous doses administered on postnatal days 0–4 to male Sprague–Dawley pups. Reduced plasma testosterone, testis weight, seminiferous cord diameter and steroidogenic acute regulatory protein expression. (A) Golden et al. (1998) and Dietrich (2010). Administration to pregnant women during weeks 7–35 of pregnancy. +*(1) No adverse effects observed +(2) cryptorchidism, decreased penis size and sperm counts.

microphallus, cryptorchidism, testicular hypoplasia, reduced sperm counts, and increased incidence of abnormal sperm (Dietrich, 2010; Golden et al., 1998) have made DES the prototype for chemical-induced TDS in humans, showing nearly identical effects in rats and humans (Toppari et al., 2010). Because no clinical trial had been conducted with DES to verify efficacy and optimize dosage, the total DES dose administered varied among clinics by more than an order of magnitude. Male reproductive tract abnormalities were significantly increased only among offspring of mothers enrolled in clinics that employed the higher dose regimens, i.e., administration of 12–18 g DES during pregnancy (Dietrich, 2010; Golden et al., 1998), or approximately 844–1266 µg/kg/day, assuming a body weight of 70 kg per pregnant woman. In contrast, no clear increase in incidence of male reproductive tract effects has been observed in offspring of mothers given lower dose regimens of DES, i.e., administration of 1.4 g DES during pregnancy (Dietrich, 2010; Golden et al., 1998; Leary et al., 1984), equivalent to 71 µg/kg/day for the first two weeks and 99 µg/kg/day during the entire pregnancy. Plotting these doses against data from studies conducted in rats (Fig. 4) demonstrates that effects on the developing male reproductive tract are observable in rats at DES doses approximately 1–2 orders of magnitude lower than those required to produce similar effects in humans. The data also indicate that the male rat reproductive tract is similarly sensitive across fetal, neonatal and adult life stages.

In a detailed analysis of species differences with respect to in utero DES-induced male reproductive tract anomalies, Hogan et al. (1987) compared the relative potency ratios for these effects in the mouse (the prototypical animal model for DES-induced reproductive tract effects) and humans. Depending on various assumptions, effects in humans occurred at doses from 1–2 orders of magnitude greater, to approximately equal those at which effects occurred in mice. Thus, both rat and mouse data challenge the seemingly arbitrary assumption that male reproductive tract malformations occur in humans at DES doses 200–500-fold less than required to produce effects in rodents (Kortenkamp and Faust, 2010). Although there could be speculation on mechanistic grounds as to why humans might be less sensitive than rodents to effects of DES but not to effects of other chemicals with potential anti-androgenic properties, such speculation would presumably be irrelevant within the context of the DA–CAOS concept wherein mechanistic similarity is not a criterion for predicting combination effects. Consequently, the comparison of human versus rat sensitivity to the effects of DES on the developing male reproductive tract appears to be relevant for the risk assessment of anti-androgens within the DA–CAOS concept, especially for phthalate esters. This comparison would seem to be an obligate exercise for using rat data to conduct a human health risk assessment, especially when the data are publicly available.

2.5.2. Human versus rat – finasteride

To further test the assumption that anti-androgenic chemicals affect the human male reproductive tract at doses lower than in the rat, data were compared for effects of finasteride, a human pharmaceutical prescribed for its anti-androgenic effects in the treatment of benign prostatic hypertrophy (Gormley et al., 1990). Finasteride was among the mixture of anti-androgens reported to synergistically induce reproductive tract abnormalities in male rats following in utero administration (Christiansen et al., 2009). Finasteride is a specific inhibitor of 5α reductase, the enzyme responsible for conversion of testosterone to dihydrotestosterone, the active androgen receptor ligand and agonist in humans and rodents. The inhibition of 5α reductase by finasteride is not mediated through DHT binding to the androgen receptor, thus mimicking hereditary 5α reductase deficiency where individuals with this deficiency present with poor prostatic growth (Gormley, 1992; Gormley et al., 1990). Finasteride has been shown to significantly reduce prostate size in humans and in several animal models. Both testosterone and dihydrotestosterone (DHT) are critical for normal male reproductive development. DHT is required for normal development of the external genitalia and prostate (Bowman et al., 2003).

Significant variability attends establishing a threshold for percent reduction in DHT in humans, as seen in Fig. 5. The lowest dose

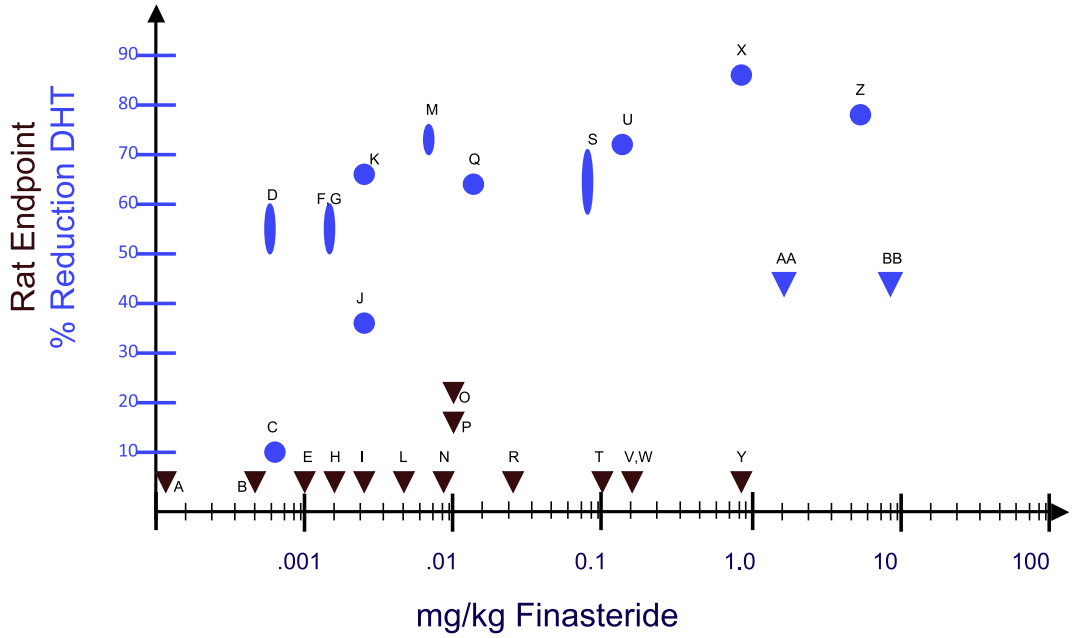

Fig. 5. Finasteride potency comparison for human clinical suppression of DHT versus rat endpoints (see Table 3).

at which a reduction of DHT was seen in men was 0.0006 mg/kg and ranged from 10% after one-day exposure to 50–60% after 14-day exposure (Gormley et al., 1990). In this study, a statistically significant percent reduction (approximately 40%) in DHT occurred in the baseline values for two treatment groups, further indicating the significant variability if this effect. The maximum suppression of DHT in serum is approximately 70% and occurs at doses > 0.007 mg/kg (Steiner, 1996). The plateau in suppression of serum DHT is shown out to 1 mg/kg in Fig. 5. Maximum suppression of DHT in the prostate is 85–90%. Finasteride has much greater affinity for the type 2 5α reductase isozyme than for the type 1, hence, the remaining DHT in the serum and prostate gland is likely to be the result of type 1 5α reductase (Bartsch et al., 2002).

Very little data were found to establish a threshold for the reduction of serum DHT in the rat. A threshold for the effect of DHT on the prostate has been demonstrated in rats, however the threshold is only apparent when the animals have been castrated, resulting in very low intra-prostatic testosterone and DHT levels (Bartsch et al., 2002). Thresholds for effects on male reproductive development in the rat using both standard developmental toxicity studies and the Hershberger assay are also shown in Fig. 5. Finasteride causes a decrease in anogenital distance in male offspring. The threshold for reversible decreased anogenital distance is 0.003 mg/kg (Clark et al., 1990) but is somewhat higher for irreversible decreased anogenital distance as determined by its presence at post-natal day 90 (Bowman et al., 2003).

Fig. 5 shows that the threshold for a clinically effective decrease (Fig. 5J) in circulating dihydrotestosterone in men occurs at approximately the same finasteride dose as produces a reduction in rat anogenital distance in the Hershberger assay (Fig. 5I). Thresholds for other effects in rats are observed at significantly higher doses, but still lower than the recommended clinical dose for treatment of benign prostatic hypertrophy in men (Fig. 5S). Taken together, the above data strongly suggest that irrespective of the high variability in both human and rat data, effects of a potent anti-androgen, finasteride, occur in human males at doses no lower than, and most likely considerably higher than are required to produce effects in the rat.

3. The Human-Relevant Potency-Threshold (HRPT)

As a conservative screening level assessment, the DA–CAOS concept may have some utility since it is reasonably simple to perform, requires only rudimentary dose–response information and demands virtually no understanding of mode of action, pharmacokinetics, or structure activity relationships. However, as demonstrated above, the DA–CAOS theory suffers an inordinate degree of uncertainty as evidenced by limitations in the studies on which it is based, contradicts fundamental tenets of pharmacology, and would predict outcomes incongruous with human clinical and epidemiological observations. The published risk assessment based on DA–CAOS magnifies those uncertainties with unnecessary conservatism regarding doses at which effects occur in humans versus rats. Thus, for any group of chemicals that warrant further analysis – for which concern might remain after conducting a DA–CAOS screening assessment – a better approach is needed that is well grounded in fundamental pharmacological principles, can be reconciled with human clinical data, and is consistent with clinical epidemiological experience.

Consequently, an improved risk assessment and prediction strategy is proposed that melds those features of the DA–CAOS concept that are tenable with requirements of the TEQ concept that are necessary to conform with fundamental pharmacological principles and to be compatible with the observed clinical and epidemiological data. The proposed approach, referred to as the Human Relevant Potency Threshold (HRPT) approach, proposes that DA be assumed for chemicals that can affect a common adverse outcome, but only at doses close to the lower limit of the observable effect range. It also proposes that DA be applied to chemicals that meet the TEQ requirements for receptor- or enzyme-mediated adverse effects and whose potency approaches that of an endogenous ligand or human pharmaceutical. In both cases, the observable effect dose or the potency of a natural ligand or human pharmaceutical should be based on human data whenever available. Below these ‘thresholds’ in either dose or potency (affinity/intrinsic activity), IA would be used for cumulative risk assessment of chemicals to which humans are exposed; i.e.,

Table 3
Human clinical data (ovals) rat experimental data (triangles). Blue points indicate % reduction in DHT.

Dose (mg/kg)	Human endpoint	Rat endpoint	Reference	
0.0001		Neg Hershberger assay	Ashby et al. (2004)	A
0.0005		Neg Hershberger assay	Ashby et al. (2004)	B
0.0006 (1 day)	10% ↓ serum DHT		Gormley et al. (1990)	C
0.0006 (14 days)	50–60% ↓ serum DHT		Gormley et al. (1990)	D
0.001		Neg Hershberger assay	Ashby et al. (2004)	E
0.0018 (1 day)	50–60% ↓ serum DHT		Gormley et al. (1990)	F
0.0018 (14 days)	50–60% ↓ serum DHT		Gormley et al. (1990)	G
0.002		Neg Hershberger assay	Ashby et al. (2004)	H
0.003		Threshold for ↓ AGD at PND 1	Clark et al. (1990)	I
0.0031 (1 day)	35% ↓ serum DHT		Gormley et al. (1990)	J
0.0031 (14 days)	65% ↓ serum DHT		Gormley et al. (1990)	K
0.005		Neg Hershberger assay	Ashby et al. (2004)	L
0.007	Reported threshold for maximal clinical suppression of DHT		Steiner (1996)	M
0.008		Threshold in Hershberger assay based on ↓ prostate and glans penis weight	Ashby et al. (2004)	N
0.01		Threshold for ↓ AGD at PND 90	Bowman et al. (2003)	O
0.01		Threshold for reversible male nipple retention	Bowman et al. (2003)	P
0.015	64% ↓ serum DHT, Suboptimal ↓ prostate volume in 6–12 mos, Suboptimal ↑ urinary flow in 6–12 mos		McConnell (1992) Gormley (1992) Gormley (1992)	Q
0.03		Threshold for permanent male nipple retention	Bowman et al. (2003)	R
0.077	Recommended clinical dose (5 mg), 59–71% ↓ serum DHT, Max ↓ prostate volume in 6–12 mos (19–26% ↓), Optimal ↑ urinary flow in 6–12 mos		Bartsch et al., 2002 Gormley (1992)	S
0.1		Threshold for hypospadias	Clark et al. (1990)	T
0.15	71% ↓ serum DHT		McConnell (1992)	U
0.2		+Hershberger assay (seminal vesicle wgt only)	Kennel et al. (2004)	V
0.2		Reported Hershberger LOELs (multiple tissues)	Owens et al. (2007)	W
0.77	85% ↓ serum DHT		McConnell (1992)	X
0.77		Lowest BMD for Hershberger assays	Owens et al. (2007)	Y
1.54	78% ↓ serum DHT, Maximum clinical dose		McConnell (1992)	Z
20		<45% ↓ serum DHT[a], 30% ↑ serum T, ↓ epididymal, prostate, SV wgt, No effect on testis wgt	Marty et al. (2001)	AA
80		45% ↓ serum DHT[a], 30% ↑ serum T, ↓ epididymal, prostate, SV wgt	Marty et al. (2001)	BB

[a] Not statistically significant.

individual RfDs would be appropriate benchmarks for risk estimation rather than hazard indices.

The HRPT approach accommodates both the DA–CAOS concept, where tenable, and the well-established TEQ concept, but improves upon each by providing a means for calibrating the assumption of DA with human data. The approach is broadly applicable whenever human data are available to support estimation of human-relevant thresholds. A conceptual diagram of the HRPT approach is presented in Fig. 6, which can be clearly contrasted with the DA–CAOS concept as depicted in Fig. 2b. Although the HRPT approach will not be possible for all chemicals due to lack of human data, the HRPT approach will be feasible in many cases, including any adverse outcome or intermediate step in the production of an adverse outcome that can be produced by a human pharmaceutical agent. For the pharmaceutical agents considered in this manuscript, DES and finasteride, clinical and rodent data were readily available in the published literature.

The steps necessary for applying the HRPT approach in cumulative (combined exposures) risk assessment include:

1. Defining the common adverse outcome or target organ effect upon which the cumulative effect from a group of chemicals is to be based. In most instances, this will be defined based on animal toxicology studies, but care should be taken to avoid effects that are arguably species specific and of questionable relevance to humans. It is important that the effect is defined specifically and that constellations of effects be avoided unless clearly and definitively related by physiological and mechanistic understanding.

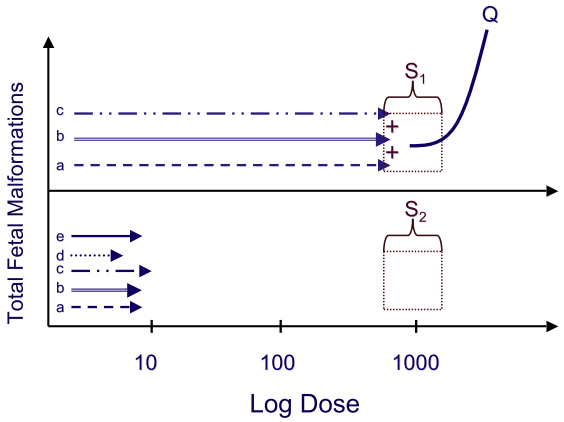

Fig. 6. HRPT conceptual model. Different line patterns a–e represent chemicals with anti-androgenic potential that produce CAOS via different modes of action. (S) indicates the dose range where mixture experiments (Table 2) have been performed relative to the dose–response curves for observable effects (Q). The HRPT approach proposes that DA be assumed for chemicals that can affect a common adverse outcome, but only at doses (S_1) that approach the lower limit of the observable effect range (Q), as depicted in the top panel. For cumulative risk assessments, IA would be used to combine chemicals to which humans are exposed that fall below these 'thresholds' (S_2, bottom panel) in either dose or potency.

2. Identifying the chemicals known to produce the common adverse outcome in the test species and determining whether data indicate DA combined effects of those chemicals. If demonstrable, identify the lowest concentrations in mixtures at which DA occurs.
3. Defining, if possible, the modes of action that can lead to the adverse outcome in the test species.
4. Identifying chemicals, including drugs, known to produce the adverse outcome in humans, including among these, chemicals or drugs known to operate by relevant modes of action that can produce the adverse outcome.
5. Identifying chemicals for which the TEQ concept is justified based on satisfying TEQ requirements (see Table 1).
6. Gathering and comparing dose–response data for the chemicals and drugs of interest in humans and the test species, whether based on end-organ toxic effects or intermediate, obligate steps in the production of toxicity. An example of the former is the comparison of the DES doses at which male reproductive tract malformations occur in humans versus rats; the latter, the comparisons based on doses of finasteride. Fortunately, because data providing direct dose sensitivity comparisons for frank adverse effects are rare, such as those for DES, comparisons based on measures of pharmacological effects will often be required.
7. Based on the comparisons of human versus test species sensitivity, potency differences between chemicals, and concentrations at which DA adverse effects are demonstrable in test species, estimating the potency differential between species, and thus the potency threshold at which DA would be a conservative but tenable assumption for humans. Defining these potency thresholds is required for TEQ-compliant chemicals as well as for broader groups based only on common adverse outcome or intermediate steps in toxicity.

4. Application of the HRPT approach to potential anti-androgens

The first three steps for applying the HRPT approach as outlined above were assumed from the NRC report (NRC, 2008) and the mixtures studies conducted on chemicals with potential anti-androgenic properties (see Table 2). Step 4 involved identification of DES as a chemical known to produce the common adverse outcome in humans, and identification of finasteride, a chemical included in the one of the subject mixtures studies (Christiansen et al., 2009) which has a defined mode of action and clearly measurable clinical endpoint in humans. Chemicals that are potent inhibitors of 5-alpha reductase, androgen receptor antagonists, and phthalate esters that interrupt Leydig and Sertoli cell development may be candidates for three separate TEQ groupings (Step 5). Step 6 has been outlined previously in this paper and is summarized in Figs. 4 and 5. Step 7 is described below using phthalate esters and finasteride as examples to illustrate how the HRPT approach is applied based on dose level and potency respectively.

For the broad grouping of chemicals with potential anti-androgenic properties, an HRPT should be set on dose level rather than potency because these chemicals do not satisfy TEQ criteria for application of DA across untested dose ranges. Based on the analysis outlined in steps 1 through 6 and the analysis presented in Section 2.4.1 and Fig. 4, a proposed dose-based HRPT for the broad grouping of phthalate esters and other chemicals with potential anti-androgenic properties can be conservatively set at doses 5-fold below rat LOAELs/NOAELs for CAOS on the developing male reproductive tract. Thus, for example, the HRPT approach would apply DA to combined human phthalate exposures that are within a factor of 5-fold lower (i.e., 20%) than rat LOAELs/NOAELs for effects on the male reproductive tract, but would apply IA to lower exposure levels, i.e., for cumulative risk assessment at current levels of human exposure to phthalates. Since the DES dose required for human in utero effects is approximately 1–2 orders of magnitude greater than required for rats (Fig. 4), an HRTP above the rat LOAELs/NOAELs may be justifiable. Thus, the proposed dose-based HRPT of 5-fold below the rat LOAELs/NOAELs for assuming DA in a cumulative risk assessment is a conservative estimate for TDS-like effects on the developing male reproductive tract of humans, and likely provides at least an additional order of magnitude conservatism. Although the DA–CAOS concept does not require a mechanistic rationale for estimating combined effects, the fact that both phthalates and DES produce effects on the developing male reproductive tract secondary to inhibition of Leydig and Sertoli cell function provides relatively high confidence that a dose-based HRPT of 5-fold below the rat NOAEL is adequately protective of human health.

Over the past decade, a number of epidemiology studies have appeared in the scientific literature that compared urinary concentrations of phthalate metabolites and developmental effects in infants, including anti-androgenic effects (e.g., cryptorchidism, anogenital distance). Those studies have a number of features in common: a proposed hypothesis based on the results of animal studies, limited populations, limited number of biomarker samples during pregnancy, and limited concurrence with animal data. Although virtually all of the studies report one or more statistically significant associations, none offer a basis for testing the HRPT approach on the potential for anti-androgenic effects in infants exposed to individual or multiple phthalates. Causal interpretation of the epidemiological findings is problematic due to several limiting factors, which are often well described by the investigators. In addition, there is a considerable amount of inconsistency among the human studies and between the animal and human studies.

Human versus rat comparisons for the potency of finasteride suggest a conservative potency-based HRPT of 1 order of magnitude below the potency of finasteride for effects on the rat male reproductive tract from androgen deficiency via inhibition of 5-alpha-reductase. This potency-based HRPT would be applied to chemicals meeting TEQ criteria for similarity of mode of action and structure–activity parameters with finasteride, and provides approximately an order of magnitude conservatism based on data indicating that humans have similar sensitivity as rats to effects of finasteride, as explained above in Section 2.4.2 and illustrated in Fig. 5. This potency-based HRPT would trigger the assumption of DA for all chemicals meeting TEQ criteria whose potency for inhibition of 5-alpha reductase is within one order of magnitude that of finasteride. Chemicals with lower potency would be assessed by IA, or by DA using an HRPT set on dose, in this instance, at doses within 5-fold below their individual NOAELs/LOAELs for CAOS on the developing male reproductive tract. Although we have not derived potency-based HRPTs specifically for androgen receptor antagonists or for inhibitors of particular enzymes in the steroidogenic pathway, these would be derived as was done for finasteride if cumulative assessments are desired for exposure to groups of chemicals that act specifically by these modes of action.

Applying the HRPT approach to the risk assessment for 15 anti-androgens results in an estimate of risk for phthalates 2 or more orders of magnitude lower than concluded in the published assessment (Kortenkamp and Faust, 2010). This is due to the fact that a conservative HRPT of 5-fold below the rat NOAEL precludes application of DA to lower human exposures, whereas Kortenkamp and Faust (2010) assumed DA for all doses and ratios based on human RfDs 200–500-fold lower than rat NOAELs. In contrast, use of the more biologically plausible HRPT approach leads to a conclusion that under current exposure conditions, phthalates should not be assessed by DA, and that few, if any of the other chemicals included in the published risk assessment (Kortenkamp and Faust, 2010)

should be assessed by DA at current exposure levels. Instead, IA should be applied to those chemicals.

5. Conclusions

Limitations in the study designs for mixtures of rodent anti-androgens (Table2), albeit imposed by practical necessity for studies in live animals, impart an undefined level of uncertainty to the extrapolation of experimental results beyond the doses and ratios tested. Additional uncertainty is introduced by the use of scored endpoints and the way these scores were combined for some analyses. A statistical analysis of precision using published data from the mixtures studies in question indicates that the model predictions may vary as much as 2-fold for the dose ratio and concentrations of chemicals tested. This variance would expand with extrapolation to untested ratios and doses. Extrapolation to lower doses is totally dependent on the model employed and because no data exist to support the extrapolation, this could dramatically increase variability and lead to erroneous conclusions. These factors reduce the confidence that can be placed in the conclusion that mixtures of anti-androgens are DA, even within the dose ranges evaluated, but even moreso at untested concentrations and ratios. No objective information allows one to conclude that DA is more accurate than IA at relevant human exposure levels. Given the questions raised here (Section 2.2) and the fact that the Rider et al. (2008) and Metzdorff et al. (2007) data, taken together, indicate that DA as well as IA predict a variety of relevant responses in the developing male reproductive tract of rats, it is unclear if not unscientific to assume DA by default based on a preference for conservatism, as claimed (Kortenkamp and Faust, 2010). Since risk management decisions often involve choosing between various options, it is impossible to make those decisions in a precautionary mode without understanding the underlying accuracy of scientific assessments. Therefore, the preference of models used to scientifically assess human health risks would be best based on data rather than on presumptions about what constitutes a precautionary decision.

The DA–CAOS theory is inconsistent with established pharmacological principles that relate affinity, intrinsic activity, and efficacy to relative potency estimation, and overlooks potential dose-dependent changes in pharmacokinetic interactions. Logical predictions of the DA–CAOS concept are also inconsistent with available clinical and epidemiological information. The incidence of TDS in humans, considered by some to be consistent with phthalate syndrome in rats, is quite low, whereas applying the DA–CAOS theory to all potential anti-androgenic drugs and chemicals would predict an epidemic of the syndrome affecting nearly the entire human population. The strong suggestion of a clinical threshold for DES-induced male reproductive tract malformations is inconsistent with the assumption of DA–CAOS, given the fact that human exposure to many anti-androgenic drugs and chemicals was significant during the DES episode. If combined exposures to chemicals that can affect a CAOS truly operate by DA irrespective of their potencies and concentrations, it is difficult to imagine how living organisms could survive in a world composed of hundreds of thousands of chemicals, all of which produce overt toxicity at some level of exposure. Indeed, the DA–CAOS theory contravenes well-established principles of pharmacological and toxicological action evidenced by mechanistic and clinical data derived from human pharmaceutical experience.

Comparison of published data for human and rat male reproductive tract sensitivities to DES and finasteride reveals that relying on rat data while ignoring the relative sensitivity of the human fetus has introduced considerable but unnecessary uncertainty and conservatism to the published human health risk assessments for anti-androgens (Kortenkamp and Faust, 2010; Benson, 2009). The magnitude of this unnecessary conservatism is at least 2 orders of magnitude, and perhaps as much as 4 orders of magnitude. Taken together, the uncertainties and conservatism introduced by applying DA to all anti-androgens at all doses, irrespective of mechanism, and ignoring human versus rat sensitivity render scientific conclusions based on the mixture studies at issue tenuous and regulatory decisions based on the risk assessment utilizing such studies arbitrary.

The DA–CAOS recommendation and the risk assessment based upon it are radical departures from past EPA practice regarding similarity criteria for applying dose addition (Table 1)(Borgert et al., 2004). Furthermore, this approach appears to ignore the entire logic of the TEQ requirements, which were developed as inclusion criteria for applying DA in cumulative risk assessments, i.e., to increase the reliability of extrapolating the DA assumption beyond empirical data. Nonetheless, because sophisticated pharmacological and toxicological knowledge is not required for its application, the DA–CAOS concept could be rationally applied, but only as a coarse screening level assessment. This is consistent with the historical use of hazard-index based approaches, which are used in screening-level baseline risk assessments to project acceptable cleanup criteria for hazardous waste sites.

If a DA–CAOS-type assessment suggests that human exposure to some groups of chemicals may exceed its conservative parameters, a more biologically based method of assessing actual risk is warranted. We have proposed such a method – the HRPT approach – to fulfill this need. The HRPT approach builds upon tenable assumptions of the DA–CAOS approach and the more biologically based TEQ approach used for decades by EPA and other regulatory bodies, but offers a means of improving the accuracy and reliability of the risk assessment by incorporating human data into the potency and combined-effect analysis. The HRPT approach is widely applicable and is feasible for any type of effect that is produced by groups of chemicals for which direct human data are available. Given the wide array of pharmacological modalities by which human pharmaceuticals act, sufficient data are available for applying the HRPT approach broadly in toxicological risk assessment.

6. Conflict of interest statement

The authors have no conflicts of interest that affect their scientific analysis or conclusions. There are no contractual relations or proprietary considerations that restrict the authors' publication or dissemination of their findings. C.J. Borgert received financial support to undertake portions of this analysis from the American Chemistry Council. The analysis and views expressed here are those of the authors and do not necessarily reflect those of the American Chemistry Council or its members.

References

Adachi, T., Koh, K.B., Tainaka, H., Matsuno, Y., Ono, Y., Sakurai, K., et al., 2004. Toxicogenomic difference between diethylstilbestrol and 17beta-estradiol in mouse testicular gene expression by neonatal exposure. Molecular Reproduction and Development 67 (1), 19–25.

Adamsson, N.A., Brokken, L.J., Paranko, J., Toppari, J., 2008. In vivo and in vitro effects of flutamide and diethylstilbestrol on fetal testicular steroidogenesis in the rat. Reproductive Toxicology 25 (1), 76–83.

Adeshina, F. and Todd, E.L., 1991. Exposure assessment of chlorinated pesticides in the environment. Journal of Environmental Science and Health: Part A: Environmental Science and Engineering (USA).

Andric, S.A., Kostic, T.S., Dragisic, S.M., Andric, N.L., Stojilkovic, S.S., Kovacevic, R.Z., 2000a. Acute effects of polychlorinated biphenyl-containing and -free transformer fluids on rat testicular steroidogenesis. Environmental Health Perspectives 108 (10), 955–959.

Andric, S.A., Kostic, T.S., Stojilkovic, S.S., Kovacevic, R.Z., 2000b. Inhibition of rat testicular androgenesis by a polychlorinated biphenyl mixture aroclor 1248. Biology of Reproduction 62 (6), 1882–1888.

Ashby, J., 2003. Problems associated with the recognition and confirmation of low-dose endocrine toxicities. Nonlinearity in Biology, Toxicology. Medicine 1 (4), 439–453.

Ashby, J., Lefevre, P.A., Tinwell, H., Odum, J., Owens, W., 2004. Testosterone-stimulated weanlings as an alternative to castrated male rats in the hershberger anti-androgen assay. Regulatory Toxicology and Pharmacology 39 (2), 229–238.

ATSDR (Agency for Toxic Substances and Disease Registry), 2001a. Guidance Manual for the Assessment of Joint Toxic Action of Chemical Mixtures. Final/ Technical Report. U.S. Department of Health and Human Services, Public Health Service Agency for Toxic Substance and Disease Registry, Atlanta, GA, USA.

ATSDR (Agency for Toxic Substances and Disease Registry), 2001b. Guidance for the Preparation of an Interaction Profile.

Axmon, A., Hagmar, L., Jönsson, B.A., 2008. Rapid decline of persistent organochlorine pollutants in serum among young swedish males. Chemosphere 70 (9), 1620–1628.

Barton, C.N., Braunberg, R.C., Friedman, L., 1993. Nonlinear statistical models for the joint action of toxins. Biometrics 49, 95–105.

Bartsch, G., Rittmaster, R.S., Klocker, H., 2002. Dihydrotestosterone and the concept of 5α-reductase inhibition in human benign prostatic hyperplasia. European Urology 37 (4), 367–380.

Benson, R., 2009. Hazard to the developing male reproductive system from cumulative exposure to phthalate esters–dibutyl phthalate, diisobutyl phthalate, butylbenzyl phthalate, diethylhexyl phthalate, dipentyl phthalate, and diisononyl phthalate. Regulatory Toxicology and Pharmacology 53 (2), 90–101.

Borgert, C.J., Price, B., Wells, C.S., Simon, G.S., 2001. Evaluating chemical interaction studies for mixture risk assessment. Human and Ecological Risk Assessment 7 (2), 259–306.

Borgert, C.J., Quill, T.F., McCarty, L.S., Mason, A.M., 2004. Can mode of action predict mixtures toxicity for risk assessment? Toxicology and Applied Pharmacology 201 (2), 85–96.

Borgert, C.J., Borgert, S.A., Findley, K.C., 2005. Synergism, antagonism, or additivity of dietary supplements: application of theory to case studies. Thrombosis Research 117 (1–2), 123–132.

Bowman, C.J., Barlow, N.J., Turner, K.J., Wallace, D.G., Foster, P.M.D., 2003. Effects of in utero exposure to finasteride on androgen-dependent reproductive development in the male rat. Toxicological Sciences 74 (2), 393–406.

Cassee, F.R., Groten, J.P., Bladeren, P.J., Feron, V.J., 1998. Toxicological evaluation and risk assessment of chemical mixtures. Critical Reviews in Toxicology 28 (1), 73–101.

CDC, 2005. Third National Report on Human Exposure to Environmental Chemicals, Atlanta, GA., NCEH Pub. No. 05-0570. July 2005.

Christiansen, S., Scholze, M., Axelstad, M., Boberg, J., Kortenkamp, A., Hass, U., 2008. Combined exposure to anti-androgens causes markedly increased frequencies of hypospadias in the rat. International Journal of Andrology 31 (2), 241–248.

Christiansen, S., Scholze, M., Dalgaard, M., Vinggaard, A.M., Axelstad, M., Kortenkamp, A., Hass, U., 2009. Synergistic disruption of external male sex organ development by a mixture of four antiandrogens. Environmental Health Perspectives 117 (12), 1839–1846.

Clark, R.L., Antonello, J.M., Grossman, S.J., et al., 1990. External genitalia abnormalities in male rats exposed in utero to finasteride, a 5 alpha-reductase inhibitor. Teratology 42 (1), 91–100.

Cook, J.C., Klinefelter, G.R., Hardisty, J.F., Sharpe, R.M., Foster, P.M., 1999. Rodent leydig cell tumorigenesis: a review of the physiology, pathology, mechanisms, and relevance to humans. Critical Reviews in Toxicology 29 (2), 169–261.

Couse, J.F., Korach, K.S., 2004. Estrogen receptor-alpha mediates the detrimental effects of neonatal diethylstilbestrol (DES) exposure in the murine reproductive tract. Toxicology 205 (1–2), 55–63.

David, R.M., 2006. Proposed mode of action for in utero effects of some phthalate esters on the developing male reproductive tract. Toxicologic Pathology 34, 209–219.

Dietrich, D.R., 2010. Courage for simplification and imperfection in the 21st century assessment of "endocrine disruption". ALTEX 27 (4), 264–278.

Efron, B., Tibshirani, R., 1993. An Introduction to the Bootstrap. Springer-Verlag, New York.

Fernandez, M.F., Olmos, B., Granada, A., López-Espinosa, M.J., Molina-Molina, J.M., Fernandez, J.M., et al., 2007. Human exposure to endocrine-disrupting chemicals and prenatal risk factors for cryptorchidism and hypospadias: a nested case-control study. Environmental Health Perspectives 115 (Suppl. 1), 8–14.

Filipiak, E., Walczak-Jedrzejowska, R., Oszukowska, E., et al., 2009. Xenoestrogens diethylstilbestrol and zearalenone negatively influence pubertal rat's testis. Folia Histochemica Et Cytobiologica, Polish Academy of Sciences. Polish Histochemical and Cytochemical Society 47 (5), S113–S120.

Fisch, H., Golden, R.J., Liberson, G.L., Hyun, G.S., Madsen, P., New, M.I., Hensle, T.W., 2001. Maternal age as a risk factor for hypospadias. Journal of Urology 165, 934–936.

Foster, P.M., 2005. Mode of action: impaired fetal Leydig cell function-effects on male reproductive development produced by certain phthalate esters. Critical Reviews in Toxicology 35 (8–9), 713–719.

Gaido, K.W., Hensley, J.B., Liu, D., et al., 2007. Fetal mouse phthalate exposure shows that Gonocyte multinucleation is not associated with decreased testicular testosterone. Toxicological Sciences 97, 491–503.

Golden, R.J., Noller, K.L., Titus-Ernstoff, L., Kaufman, R.H., Mittendorf, R., Stillman, R., Reese, E.A., 1998. Environmental endocrine modulators and human health: an assessment of the biological evidence. Critical Reviews in Toxicology 28 (2), 109–227.

Gormley, G.J., 1992. Chemoprevention strategies for prostate cancer: the role of 5 alpha-reductase inhibitors. Journal of Cellular Biochemistry, Supplement 16H, 113–117.

Gormley, G.J., Stoner, E., Rittmaster, R.S., et al., 1990. Effects of finasteride (MK-906), a 5 alpha-reductase inhibitor, on circulating androgens in male volunteers. The Journal of Clinical Endocrinology and Metabolism 70 (4), 1136–1141.

Goyal, H.O., Braden, T.D., Mansour, M., et al., 2001. Diethylstilbestrol-treated adult rats with altered epididymal sperm numbers and sperm motility parameters, but without alterations in sperm production and sperm morphology. Biology of Reproduction 64 (3), 927–934.

Goyal, H.O., Braden, T.D., Williams, C.S., et al., 2004. Abnormal morphology of the penis in male rats exposed neonatally to diethylstilbestrol is associated with altered profile of estrogen receptor-alpha protein, but not of androgen receptor protein: a developmental and immunocytochemical study. Biology of Reproduction 70 (5), 1504–1517.

Goyal, H.O., Braden, T.D., Williams, C.S., et al., 2005. Permanent induction of morphological abnormalities in the penis and penile skeletal muscles in adult rats treated neonatally with diethylstilbestrol or estradiol valerate: a dose-response study. Journal of Andrology 26 (1), 32–43.

Goyal, H.O., Braden, T.D., Williams, C.S., Williams, J.W., 2009. Estrogen-Induced developmental disorders of the rat penis involve both estrogen receptor (ESR)- and androgen receptor (AR)-mediated pathways. Biology of Reproduction 81 (3), 507–516.

Gray, L.E., Kelce, W.R., Monosson, E., Ostby, J.S., Birnbaum, L.S., 1995. Exposure to TCDD during development permanently alters reproductive function in male long evans rats and hamsters: reduced ejaculated and epididymal sperm numbers and sex accessory gland weights in offspring with normal androgenic status. Toxicology and Applied Pharmacology 131 (1), 108–118.

Gray, L.E., Ostby, J.S., Kelce, W.R., 1997. A dose-response analysis of the reproductive effects of a single gestational dose of 2, 3, 7, 8-tetrachlorodibenzo-p-dioxin in male long evans hooded rat offspring. Toxicology and Applied Pharmacology 146 (1), 11–20.

Gray, L.E., Wolf, C., Lambright, C., et al., 1999. Administration of potentially antiandrogenic pesticides (procymidone, linuron, iprodione, chlozolinate, p,p′–DDE, and ketoconazole) and toxic substances (dibutyl– and diethylhexyl phthalate, PCB 169, and ethane dimethane sulphonate) during sexual differentiation produces diverse profiles of reproductive malformations in the male rat. Toxicology and Industrial Health 15 (1–2), 94–118.

Guyot, R., Odet, F., Leduque, P., Forest, M.G., Le Magueresse-Battistoni, B., 2004. Diethylstilbestrol inhibits the expression of the steroidogenic acute regulatory protein in mouse fetal testis. Molecular and Cellular Endocrinology 220 (1–2), 67–75.

Haavisto, T., Nurmela, K., Pohjanvirta, R., et al., 2001. Prenatal testosterone and luteinizing hormone levels in male rats exposed during pregnancy to 2, 3, 7, 8-tetrachlorodibenzo-p-dioxin and diethylstilbestrol. Molecular and Cellular Endocrinology 178 (1–2), 169–179.

Haavisto, T.E., Adamsson, N.A., Myllymaki, S.A., et al., 2003. Effects of 4-tert-octylphenol, 4-tert-butylphenol, and diethylstilbestrol on prenatal testosterone surge in the rat. Reproductive Toxicology 17 (5), 593–605.

Hallmark, N., Walker, M., McKinnell, C., et al., 2007. Effects of monobutyl and di(n-butyl) phthalate in vitro on steroidogenesis and Leydig cell aggregation in fetal testis explants from the rat: comparison with effects in vivo in the fetal rat and neonatal marmoset and in vitro in the human. Environmental Health Perspectives 115, 390–396.

Haschek W.M., Rousseaux C.G., Wallig M.A. (eds.), 2010. Nomenclature: terminology for Morphological Alterations, Chapter 4. Fundamentals of Toxicologic Pathology, 2nd ed. London.

Hass, U., Scholze, M., Christiansen, S., et al., 2007. Combined exposure to anti-androgens exacerbates disruption of sexual differentiation in the rat. Environmental Health Perspectives 115 (Suppl. 1), 122–128.

Hays, S.M., Aylward, L.L., 2001. Temporal trends in body-burden suggest that dioxin exposure in the general population have declined significantly. Organohalogen Compounds 52, 214.

Hogan, M.D., Newbold, R.R., McLachlan, J.A., 1987. Extrapolation of teratogenic responses observed in laboratory animals to humans: DES as an illustrative example. In: McLachlan, J.A., Pratt, R.M., Markert, C.L. (Eds.), Developmental Toxicology: Mechanisms and Risks, Banbury Report 26. Cold Spring Harbor Laboratory, Cold Spring Harbor, pp. 257–269.

Hotchkiss, A.K., Parks-Saldutti, L.G., Ostby, J.S., et al., 2004. A mixture of the "antiandrogens" linuron and butyl benzyl phthalate alters sexual differentiation of the male rat in a cumulative fashion. Biology of Reproduction 71 (6), 1852–1861.

Hovinga, M.E., Sowers, M., Humphrey, H.E.B., 1992. Historical changes in serum PCB and DDT levels in an environmentally-exposed cohort. Archives of Environmental Contamination and Toxicology 22 (4), 362–366.

Howdeshell, K.L., Furr, J., Lambright, C.R., et al., 2007. Cumulative effects of dibutyl phthalate and diethylhexyl phthalate on male rat reproductive tract development: altered fetal steroid hormones and genes. Toxicological Sciences 99 (1), 190–202.

Howdeshell, K.L., Rider, C.V., Wilson, V.S., et al., 2008a. Mechanisms of action of phthalate esters, individually and in combination, to induce abnormal reproductive development in male laboratory rats. Environmental Research 108 (2), 168–176.

Howdeshell, K.L., Wilson, V.S., Furr, J., et al., 2008b. A mixture of five phthalate esters inhibits fetal testicular testosterone production in the Sprague-Dawley rat in a cumulative, dose-additive manner. Toxicological Sciences 105 (1), 153–165.

Ikeda, Y., Tanaka, H., Esaki, M., 2008. Effects of gestational diethylstilbestrol treatment on male and female gonads during early embryonic development. Endocrinology 149 (8), 3970–3979.

Johnson, K.J., McDowell, E.N., Viereck, M.P., et al., 2011. Species-specific dibutyl phthalate fetal testis endocrine disruption correlates with inhibition of SREBP2-dependent gene expression pathways. Toxicological Sciences 120 (2), 460–474.

Kelce, W.R., Stone, C.R., Laws, S.C., Gray, L.E., Kemppainen, J.A., Wilson, E.M., 1995. Persistent DDT metabolite p,p′-DDE is a potent androgen receptor antagonist. Nature 375 (6532), 581–585.

Kennel, P.F., Pallen, C.T., Bars, R.G., 2004. Evaluation of the rodent hershberger assay using three reference endocrine disrupters (androgen and antiandrogens). Reproductive Toxicology 18 (1), 63–73.

Khan, S.A., Ball, R.B., Hendry, W.J., 1998. Effects of neonatal administration of diethylstilbestrol in male hamsters: disruption of reproductive function in adults after apparently normal pubertal development. Biology of Reproduction 58 (1), 137–142.

Klinger, M.M., MacCarter, G.D., Boozer, C.N., 1996. Body weight and composition in the sprague dawley rat: Comparison of three outbred sources. Laboratory Animal Science 46 (1), 67–70.

Kortenkamp, A., Faust, M., 2010. Combined exposures to anti–androgenic chemicals: Steps towards cumulative risk assessment. International Journal of Andrology 33 (2), 463–474.

Krishnan, K., Brodeur, J., 1994. Toxic interactions among environmental pollutants: Corroborating laboratory observations with human experience. Environmental Health Perspectives 102 (Suppl. 9), 11–17.

Kristensen, D.M., Hass, U., Lesné, L., Lottrup, G., Jacobsen, P.R., Desdoits-Lethimonier, C., et al., 2011. Intrauterine exposure to mild analgesics is a risk factor for development of male reproductive disorders in human and rat. Human Reproduction (Oxford, England) 26 (1), 235–244.

Lambrot, R., Muczynski, V., Lécureuil, C., et al., 2009. Phthalates impair germ cell development in the human fetal testis in vitro without change in testosterone production. Environmental Health Perspectives 117 (1), 32–37.

Laska, E.M., Meisner, M., Siegel, C., 1994. Simple designs and model-free tests for synergy. Biometrics 50, 834–841.

Lassurguere, J., Livera, G., Habert, R., et al., 2003. Time- and dose-related effects of estradiol and diethylstilbestrol on the morphology and function of the fetal rat testis in culture. Toxicological Sciences 73 (1), 160–169.

Leary, F.J., Resseguie, L.J., Kurland, L.T., O'Brien, P.C., Emslander, R.F., Noller, K.L., 1984. Males exposed in utero to diethylstilbestrol. J. Amer. Med. Assoc. 252, 2984.

Marty, M.S., Crissman, J.W., Carney, E.W., 2001. Evaluation of the male pubertal onset assay to detect testosterone and steroid biosynthesis inhibitors in CD rats. Toxicological Sciences 60 (2), 285–295.

Mathews, E., Braden, T.D., Williams, C.S., et al., 2009. Mal-Development of the penis and loss of fertility in male rats treated neonatally with female contraceptive 17alpha-ethinyl estradiol: a dose-response study and a comparative study with a known estrogenic teratogen diethylstilbestrol. Toxicological Sciences 112 (2), 331–343.

McConnell, J.D., 1992. The role of dihydrotestosterone in benign prostatic hyperplasia. Current Opinion in Urology 2 (1), 18.

McKinnell, C., Atanassova, N., Williams, K., et al., 2001. Suppression of androgen action and the induction of gross abnormalities of the reproductive tract in male rats treated neonatally with diethylstilbestrol. Journal of Andrology 22 (2), 323–338.

Metzdorff, S.B., Dalgaard, M., Christiansen, S., et al., 2007. Dysgenesis and histological changes of genitals and perturbations of gene expression in male rats after in utero exposure to antiandrogen mixtures. Toxicological Sciences 98 (1), 87–98.

Mikkilä, T.F., Toppari, J., Paranko, J., 2006. Effects of neonatal exposure to 4-tert-octylphenol, diethylstilbestrol, and flutamide on steroidogenesis in infantile rat testis. Toxicological Sciences 91 (2), 456–466.

Mileson, B.E., Chambers, J.E., Chen, W.L., Dettbarn, W., Ehrich, M., Eldefrawi, A.T., et al., 1998. Common mechanism of toxicity: A case study of organophosphorus pesticides. Toxicological Sciences : An Official Journal of the Society of Toxicology 41 (1), 8–20.

NRC (National Research Council), 2008. Phthalates and Cumulative Risk Assessment: The Task Ahead. The National Acadamies Press, Washington, DC.

Owens, W., Gray, L.E., Zeiger, E., et al., 2007. The OECD program to validate the rat hershberger bioassay to screen compounds for in vivo androgen and antiandrogen responses: phase 2 dose–response studies. Environmental Health Perspectives 115 (5), 671–678.

Petreas M., She J., Visita P., Winkler J., McKinney M., Brown F.R., Dhaliwal J., Denison G., Mok M., 2001. Trends in persistent contaminants in California biota, Symposia Papers Presented Before the Division of Environmental Chemistry, American Chemical Society, San Diego, CA, April 1–5, 2001.

Price, B., Borgert, C.J., Wells, C.S., Simon, G.S., 2002. Assessing toxicity of mixtures: the search for economical study designs. Human and Ecological Risk Assessment: An International Journal 8 (2), 305–326.

Pullen, A.H., 1976. A parametric analysis of the growing CFHB (wistar) rat. Journal of Anatomy 121 (Pt 2), 371–383.

Rider, C.V., Furr, J., Wilson, V.S., et al., 2008. A mixture of seven antiandrogens induces reproductive malformations in rats. International Journal of Andrology 31 (2), 249–262.

Rider, C.V., Wilson, V.S., Howdeshell, K.L., et al., 2009. Cumulative effects of in utero administration of mixtures of "antiandrogens" on male rat reproductive development. Toxicologic Pathology 37 (1), 100–113.

Rivas, A., McKinnell, C., Fisher, J.S., Atanassova, N., Williams, K., Sharpe, R.M., 2003. Neonatal coadministration of testosterone with diethylstilbestrol prevents diethylstilbestrol induction of most reproductive tract abnormalities in male rats. Journal of Andrology 24 (4), 557–567.

Safe, S.H., 1990. Polychlorinated biphenyls (pcbs), dibenzo-p-dioxins (pcdds), dibenzofurans (pcdfs), and related compounds: environmental and mechanistic considerations which support the development of toxic equivalency factors (tefs). Critical Reviews in Toxicology 21 (1), 51–88.

Safe, S.H., 1998. Hazard and risk assessment of chemical mixtures using the toxic equivalency factor approach. Environmental Health Perspectives 106 (Suppl. 4), 1051–1058.

Scott, H.M., Mason, J.I., Sharpe, R.M., 2009. Steroidogenesis in the fetal testis and its susceptibility to disruption by exogenous compounds. Endocrine Reviews 30 (7), 883–925.

Sharpe, R.M., 2003. The 'oestrogen hypothesis'- where do we stand now? International Journal of Andrology 26 (1), 2–15.

Skakkebaek, N.E., Rajpert-De Meyts, E., Main, K.M., 2001. Testicular dysgenesis syndrome: an increasingly common developmental disorder with environmental aspects: Opinion. Human Reproduction 16 (5), 972.

Slikker, W., Andersen, M.E., Bogdanffy, M.S., Bus, J.S., Cohen, S.D., Conolly, R.B., et al., 2004. Dose-Dependent transitions in mechanisms of toxicity. Toxicology and Applied Pharmacology 201 (3), 203–225.

Steiner, J.F., 1996. Clinical pharmacokinetics and pharmacodynamics of finasteride. Clinical pharmacokinetics 30 (1), 16–27.

Tallarida, R.J., 2006. An overview of drug combination analysis with isobolograms. The Journal of Pharmacology and Experimental Therapeutics 319 (1), 1–7.

Tallarida, R.J., 2007. Interactions between drugs and occupied receptors. Pharmacology & Therapeutics 113 (1), 197–209.

Thorup, J., McLachlan, R., Cortes, D., Nation, T.R., Balic, A., Southwell, B.R., et al., 2010. What is new in cryptorchidism and hypospadias–a critical review on the testicular dysgenesis hypothesis. Journal of Pediatric Surgery 45 (10), 2074–2086.

Toppari, J., Virtanen, H.E., Main, K.M., Skakkebaek, N.E., 2010. Cryptorchidism and hypospadias as a sign of testicular dysgenesis syndrome (TDS): environmental connection. Birth Defects Research. Part A. Clinical and Molecular Teratology 88 (10), 910–919.

USDHHS (U.S. Department of Health and Human Services, Public Health Service Agency for Toxic Substance and Disease Registry), 2001. Guidance Manual for the Assessment of Joint Toxic Action of Chemical Mixtures. Final/Technical Report. Atlanta, GA, USA.

USEPA (United States Environmental Protection Agency), 1986. Guidelines for the Health Risk Assessment of Chemical Mixtures U.S. Environmental Protection Agency, Washington, DC.

USEPA (United States Environmental Protection Agency), 1989. Office of Emergency and Remedial Response. Risk assessment guidance for Superfund. Volume 1 Human health evaluation manual (Part A). Interim final U.S. Environmental Protection Agency, Washington, DC.

USEPA (United States Environmental Protection Agency), 1999. Guidance for Identifying Pesticide Chemicals That Have a Common Mechanism of Toxicity U.S. Environmental Protection Agency, Washington, DC.

USEPA (United States Environmental Protection Agency), 2000. Supplementary Guidance for Conducting Health Risk Assessment of Chemical Mixtures U.S. Environmental Protection Agency, Washington, DC.

USEPA, 2006. An inventory of sources and environmental releases of dioxin-like compounds in the United States for the years 1987, 1995, and 2000. National Center for Environmental Assessment, Washington, DC, EPA/600/P-03/002F.

Warita, K., Mitsuhashi, T., Sugawara, T., Tabuchi, Y., Tanida, T., Wang, Z.Y., et al., 2010. Direct effects of diethylstilbestrol on the gene expression of the cholesterol side-chain cleavage enzyme (p450scc) in testicular leydig cells. Life Sciences 87 (9–10), 281–285.