



EPA Public Access

Author manuscript

Reprod Toxicol. Author manuscript; available in PMC 2019 October 01.

About author manuscripts

Submit a manuscript

Published in final edited form as:

Reprod Toxicol. 2018 October ; 81: 259–271. doi:10.1016/j.reprotox.2018.08.016.

DEVELOPMENT OF A CURATED HERSHBERGER DATABASE

P Browne^a, NC Kleinstreuer^b, P Ceger^c, C Deisenroth^d, N Baker^e, K Markey^f, RS Thomas^d,
RJ Judson^d, W Casey^b

^aOECD/EHS Paris, France

^bNIH/NIEHS/DNTP/NICEATM, RTP, NC, USA

^cILS, RTP, NC, USA

^dEPA/ORD/NCCT, RTP, NC, USA

^eLeidos, RTP, NC, USA

^fEPA/OSCP, Washington, DC, USA

Abstract

A systematic literature review was conducted to identify Hershberger bioassays for ~3200 chemicals including those used to validate the OECD/US EPA guideline assay, US EPA's chemicals screened for endocrine activity, and the library of chemicals run in US EPA's ToxCast in vitro assays. For 134 chemicals that met pre-defined criteria, experimental results were extracted into a database used to characterize uncertainty in results and evaluate the concordance of the Hershberger assay with other in vivo rodent studies that measure androgen-responsive endpoints. Of 25 chemicals tested in >1 Hershberger study, 28% had disagreements between studies (i.e. 1 positive and 1 negative study), and of the 65 chemicals tested in Hershberger studies and other in vivo studies with androgen-responsive endpoints, 43% indicated disagreements, though in some cases these may be explained by differences in study designs or physiology of the animal model. Ultimately, 49 chemicals were identified with reproducible androgen pathway responses confirmed in 2 in vivo rodent studies that could be considered reference chemicals useful for validating alternative methods.

Keywords

Hershberger; androgenic; anti-androgenic; endocrine disruption; EDSP; systematic review; reference chemical; male reproduction

I. Introduction

The rodent Hershberger assay development began in the 1930s and was standardized in the 1950s (Hershberger et al. 1953) as an in vivo bioassay to investigate potential androgenic and anti-androgenic effects of chemicals, including chemicals that inhibit key steps in

Disclaimer: The opinions expressed and arguments employed herein are those of the authors, and do not necessarily reflect the official views of the OECD or of the governments of its member countries, nor the official policy of any federal agency.

steroidogenesis, such as the enzyme 5 α -reductase. The Hershberger assay was proposed as an EDSP Tier 1 screening assay in 1998 and following a multi-phase validation coordinated by the Organization for Economic Cooperation and Development (OECD) was finalized as both an US EPA Test Guideline (US EPA OPPTS 890.1400, US EPA 2009a) and OECD Test Guideline (OECD TG 441, OECD 2009a). The EDSP Tier 1/OECD Hershberger test guideline uses a castrated male rat model. Rats are castrated around postnatal day (PND) 42 and allowed a post-surgical recovery period of at least seven days in order for endogenous testosterone levels to decrease. Results of the assay are based on changes in weights of five androgen-dependent accessory sex tissues (ASTs) measured in Hershberger assay; the ventral prostate (VP), seminal vesicle (SV) (plus fluids and coagulating glands), levator ani-bulbocavernosus (LABC) muscle, paired Cowper's glands (COW) and the glans penis (GP).

Animals are dosed once daily by oral gavage or subcutaneous injection, with dosing route determined by relevance to human exposures - gavage simulates human oral exposure, whereas subcutaneous exposure simulates other routes that bypass first-pass hepatic metabolism. Following exposure to the test chemical alone (androgenic test phase) or co-administration of test chemical with a potent androgen (anti-androgenic test phase), animals are killed within 24 hours of the last dose and ASTs are excised. Weights of ASTs are compared to respective controls, and significant changes in the mean weights of two or more tissues indicate treatment-related effects. Terminal body weight and body weight gain are also measured; liver and kidney weights and peripheral hormone levels are optional endpoints.

Due to animal welfare concerns regarding the castration procedure, a limited OECD validation was conducted using the gonad-intact stimulated male weanling rat (PND 21/22) measuring change in weight of the same ASTs (excluding the glans penis which cannot be reliably excised in immature animals prior to preputial separation), plus the testes and epididymides (Ashby and Lefevre 2000a, Ashby et al. 2004). Results indicated the stimulated weanling model was responsive to androgenic and anti-androgenic chemicals, though greater doses of testosterone propionate (TP) were required to increase AST wts (1.0 mg/kg bw/d in the weanling versus 0.2–0.4 mg/kg bw/d in the peri-pubertal rat), and the relative increase in tissue weights were smaller compared to the responses in castrated peri-pubertal male rats (OECD 2008a, OECD 2009c). Weanling AST were also less responsive (i.e. AST weights decreased by less) in animals co-treated with TP and the androgen receptor (AR) antagonist, flutamide, perhaps due to compensatory mechanisms mediated by the intact hypothalamic-pituitary-gonadal (HPG) axis in the immature animal model (OECD 2009b). In addition, AST weights in intact adult male rats following exposure to potential anti-androgens were also evaluated (Ashby and Lefevre 2000b). Comparisons of results from the three models indicated postpubertal castrated rats were maximally sensitive for screening (anti)androgenic chemicals and this animal model was selected for the test guideline (OECD 2009a).

While the OECD validation demonstrated the utility of the Hershberger assay for detecting (anti)androgenic substances, the *in vivo* assay provides somewhat limited information and requires time, use of animals and is relatively expensive. There are tens of thousands of chemicals yet to be screened for endocrine effects (e.g. US EPA 2012). Given the resources

associated with in vivo screening methods and recent technological developments in the field of toxicology, there has been a paradigm shift towards using alternatives to animal testing for chemical hazard identification. New computational and high throughput in vitro methods are being used for initial screening and prioritization, and in some cases, novel methods may be adequate to replace animal-based test methods (e.g. Ezendam et al. 2016; Browne et al. 2015; US EPA 2015a).

Important challenges for validating any new alternative method are characterizing the variability in the reference animal method, and identifying a set of chemicals with reproducible results that can be used to evaluate performance of alternative methods. To identify high quality data for chemicals with in vivo estrogenic effects in the uterotrophic bioassay, Kleinstreuer et al. (2017) meticulously catalogued results of over 2600 uterotrophic assays published in peer-reviewed scientific literature. From the literature search, a database was created that included methodological details and study results for hundreds of chemicals, many of which were tested in multiple uterotrophic assays. Experimental designs of the identified studies were compared to the guideline method (US EPA 2009b), with the expectation that any chemical tested in a published study following an experimental design sufficiently similar to the guideline method would result in the same response (or lack of response). An additional level of scrutiny was added for candidate reference chemicals, requiring independent verification of “guideline-like” uterotrophic results in at least two laboratories (Browne et al. 2015). Interestingly, substantial variability was observed in the uterine weight response to the same chemical tested following what was rigidly defined as a “guideline” method, potentially due to factors such as differences in rodent strain/species, age of animals, reproductively immature versus ovariectomized animal models, duration of exposure, and route of chemical exposure (Kleinstreuer et al. 2017). This approach for consolidating data from existing animal studies can be adopted to characterize the reproducibility of other methods used in current practice, and to identify candidate reference chemicals for in vivo endpoints. Such reference chemicals can ultimately be used for performance-based validation of new methods, reducing or obviating the need for additional animal testing to validate non-animal alternatives and expediting the validation process to better match the rate at which new tools become available. However, making use of existing in vivo studies requires clearly defined literature search terms, careful curation of resulting data, and transparent criteria for data inclusion/exclusion.

In this manuscript, we describe a Hershberger assay database compiling study results of chemicals tested in Hershberger bioassays that were conducted as part of the test guideline method development and validation, studies submitted to US EPA in response to EDSP Tier 1 test orders, and studies identified from a systematic literature review. Details on test chemical, experimental design, and experimental results were recorded, and studies under consideration were required to meet a priori criteria to ensure quality of results. The resulting database was used to characterize the variability of the Hershberger bioassay and to examine concordance of Hershberger results with other in vivo tests measuring androgen-responsive endpoints, and to then identify a set of chemicals with reproducible in vivo androgen pathway activity. This database is intended to be a scientific resource for a variety of purposes, including performance-based validation of alternatives to in vivo testing. An example of performance-based validation is demonstrated in the companion manuscript

(Kleinstreuer et al. in prep), where reference chemicals derived from this analysis are compared to the results of the US EPA ToxCast AR model (Kleinstreuer et al. 2017).

2. Methods

2.1. Literature Curation

A comprehensive literature search was conducted to identify Hershberger studies. The US National Center for Biotechnology Information's PubMed (<https://www.ncbi.nlm.nih.gov/pubmed>) and Web of Science databases (www.webofknowledge.com) were searched for the term "Hershberger", with the term specifically excluded from author fields. Citations were extracted into a master record, and duplicate references, review articles, abstracts without associated manuscripts (e.g. conference presentations), and articles that were not in English were removed.

The resulting citations were linked to a chemical name, Medical Subject Heading (MeSH) term, and/or Chemical Abstracts Service Registry Number (CASRN) for ~3200 discrete chemicals including chemicals run in the US EPA's ToxCast AR assays (<http://epa.gov/comptox/toxcast/data.htm>), chemicals used by US EPA and OECD to develop and validate the Hershberger test guideline, EDSP List 1/Tier 1 chemicals (US EPA 2015b), and chemicals identified as potential in vitro AR reference chemicals (Kleinstreuer et al. 2017; Supplemental Table 1). Efforts were taken to identify the same chemical structure referred to by different names. All test chemical names and CASRNs reported in publications were compared against US EPA's Substance Registry System (<http://www.epa.gov/srs/>), US EPA's Aggregated Computational Toxicology Resource (ACToR; <http://actor.epa.gov/actor/faces/ACToRHome.jsp>), NIH/National Library of Medicine ChemIDplus (<http://chem.sis.nlm.nih.gov/chemidplus/chemidlite.jsp>), NIH's PubChem (<http://pubchem.ncbi.nlm.nih.gov/>), and the Royal Society of Chemistry's ChemSpider (<http://www.chemspider.com/>) databases. Identity of chemical names and CASRNs reported from the literature that matched EPA's Substance Registry System were considered confirmed. In other cases, reported chemical names were only considered to have a confirmed structure if they were identified by the same synonym or CASRN in at least two other databases identified above.

The full text of each publication was reviewed for relevance and when appropriate, study details were extracted into a database. Results from Hershberger assays for the 52 EDSP List 1 chemicals were also extracted. In many cases, a single publication included experiments using several chemicals and/or study designs. Details of each chemical/experiment/publication combination (i.e. "study") were recorded separately and information such as chemical name, purity, source, CASRN; rat strain, age, intact/castrated, post-surgical recovery duration, number of animals per treatment group; dosing duration, route of test chemical administration, dose levels; positive and negative controls for androgenic and anti-androgenic mode of the assays; ASTs measured; and significant effects were recorded (Supplemental Table 2). Study details that did not conform to systematic data collection but were otherwise relevant were collected in various "comments" fields. In addition, publications that included in vitro, toxicokinetic, or additional in vivo studies were noted. Significant increases in liver weight and decreases in body weight reported for treatment

groups compared to the appropriate controls were considered potential indicators of systemic toxicity (Supplemental Table 2), though liver weights were not always reported. Significant decreases in AST weight that occurred at doses coincident with 10% decrease in body weight or 25% increase in liver weight were noted (Supplemental Table 2).

2.2. Criteria for Study Inclusion

In order to examine consistency of chemical effects in multiple Hershberger studies, we defined criteria for including/excluding studies from this analysis based on their adherence to guideline study protocols. In contrast to the similar undertaking examining increases in uterine weights following test chemical administration in published uterotrophic studies (Kleinstreuer et al. 2016), the Hershberger assay is a considerably more complicated experimental design. The OECD/US EPA Hershberger guideline includes five AST weights, several optional organ weights, and has both androgenic and anti-androgenic portions. The final guideline Hershberger protocol treats castrated peri-pubertal male rats with test chemicals for 10 consecutive days, allows at least seven days of post-surgical recovery prior to dosing, includes six or more rats for each dose level, measures weights of five ASTs (VP, SV, LABC, CG, and GP), and conducts the necropsy 18–36 h after the last dose (Table 1). At least two dose levels are included in the androgenic mode of the assay and at least three dose levels are included in the anti-androgenic mode. The anti-androgenic mode includes coadministration of test chemical with 0.2 to 0.4 mg/kg bw/d TP, and appropriate positive and negative control groups for both the androgenic and anti-androgenic modes are required (Table 1). A significant effect of a chemical tested in the androgenic mode of the assay is an increase in the mean weight of two or more ASTs compared to vehicle controls (i.e. a “positive” androgenic effect). Similarly, a chemical is considered to have a significant effect in the anti-androgenic mode if coadministration of test chemical plus reference androgen results in significantly decreased weights of two or more ASTs relative to reference androgen controls (i.e. a “positive” anti-androgenic effect). Thus, the lowest observed effect level (LOEL) is the chemical dose that causes a significant change in at least two ASTs. The no observed effect level (NOEL) is also noted for the chemical dose at which no effect occurred (or only occurred in one AST) or the highest concentration tested where fewer than two AST weights were affected. The LOEL and NOEL are reported in mg/kg bw/d (Supplemental Table 2).

Similar to the evaluation of chemicals tested in uterotrophic studies (Kleinstreuer et al. 2017; Browne et al. 2015), we developed a priori criteria for including/excluding Hershberger data from this analysis to determine consistency of responses and identify candidate *in vivo* reference chemicals. Proposed criteria were reviewed with the Reference Chemical Working Group, convened by the Interagency Coordinating Committee on the Validation of Alternative Methods (ICCVAM; <https://ntp.niehs.nih.gov/pubhealth/evalatm/test-method-evaluations/refchem/index.html>), specifically to identify candidate reference chemicals for novel toxicological method validation. The Hershberger guideline protocol was optimized for sensitivity following several rounds of validation through the OECD, during which high quality data were generated (OECD 2008b). In the interest of including all available high quality data, the group agreed that results from chemicals tested in Hershberger protocols demonstrated to be reproducible and reliable through OECD test method optimization and

validation, but less sensitive than the final guideline method, should be considered in the circumstance where positive results were observed. The underlying logic was that if a test chemical resulted in a positive result (i.e. AST weights were significantly different from control animals) using a less sensitive protocol, then the same chemical would likely have a positive effect when tested in a more sensitive study design or animal model. Candidate negative reference chemicals (i.e. chemicals that did not produce a significant change in at least two AST weights compared to controls) were only considered from study designs that more closely adhered to the final guideline protocol. Criteria for the Hershberger test guideline and those used to identify chemicals with positive and negative results are summarized in Table 1.

2.3. Evaluation of other in vivo data

Though there are differences in the specificity and sensitivity for endocrine screening assays that include androgen responsive endpoints, we were interested in identifying chemicals with reproducible in vivo (anti)androgenic effects. Results from other male rodent assays designed to screen for (anti)androgens were compared to Hershberger data. The EDSP Tier 1 screening battery requires both a Hershberger assay and a male pubertal rat assay (US EPA 2009c), though the requirement for either assay may be satisfied by “other scientifically relevant information” (US EPA 2009d). Results of pubertal male assays conducted in response to List 1/Tier 1 test orders, as part of the US EPA guideline validation (US EPA 2007), and pubertal studies identified from a literature search for the chemicals for which we also identified Hershberger studies, were reviewed. Other in vivo male rat study designs with androgen-responsive endpoints (e.g. 28 day repeated dose oral toxicity, 90 day repeated dose oral toxicity, or extended one generation studies in rats; Supplemental Table 2) were included in the analyses if a different male rat study design was accepted by the US EPA as OSRI in lieu of a male pubertal study order for a List 1 chemical, or if the results of the other in vivo study were reported in the same publication as the Hershberger study results. The results of male pubertal assays (or other experimental protocols as defined above) were briefly summarized, and details such as the animal model or test guideline, age of animals, dosing route, NOEL, LOEL, effect, and comments were recorded (Supplemental Table 2).

To derive a list of chemicals with reproducible effects on the androgen pathway in vivo, we compared results for chemicals tested in Hershberger studies with responses reported from other in vivo studies with androgen-responsive endpoints. We considered chemicals with confirmed androgen pathway effects in more than one study (that met the criteria for inclusion defined for Hershberger studies in Table 1) according to the following criteria:

1. Positive: 2 Hershberger positive OR 1 Hershberger + 1 other in vivo (both positive), and greater number of positives than negatives
2. Negative: 2 negative Hershberger (with no positives) OR 1 negative Hershberger + 1 other negative in vivo (with no positives)

2.4 Literature search for chemicals that interact with 5 α -reductase

An independent comprehensive literature search was conducted to identify a reference list for reported 5 α -reductase inhibitors. Articles were extracted from the US National Center

for Biotechnology Information's PubMed database (<https://www.ncbi.nlm.nih.gov/pubmed>) using MeSH terms for "5-alpha Reductase Inhibitors" (MeSH uid D058891). For initial processing, all identified chemical names were extracted and filtered to remove family names that did not correspond to a unique compound. The remaining records were manually curated for relevance to 5 α -reductase inhibition, resulting in 997 citations for 161 unique compounds. The compound list was cross-referenced against the National Library of Medicine Pharmaceutical Action file for "5-alpha reductase inhibitors" (<https://www.ncbi.nlm.nih.gov/mesh/82058891>) to identify chemicals not present in the initial search. The final list contained 1001 chemical/publication pairs resulting in 849 unique citations for 165 chemical compounds (Supplemental Table 3). CASRNs were assigned for all compounds and related synonyms, where available, by searching on the US EPA Chemistry Dashboard (<https://comptox.epa.gov/dashboard>), NIH's PubChem (<http://pubchem.ncbi.nlm.nih.gov/>), and the Royal Society of Chemistry's ChemSpider (<http://www.chemspider.com/>) databases. Compounds were rank-ordered by citation frequency to identify potential reference compounds (Supplemental Table 4). The resulting list of compounds was cross-referenced to the Hershberger database (Supplemental Table 5).

3. Results

From the systematic literature review for Hershberger assays, we identified 178 publications (Figure 1). For seven of the EDSP List 1 chemicals, the Tier 1 requirement was satisfied by an existing published Hershberger study, which was usually returned in the literature search. In these instances, care was taken to eliminate duplications. The resulting compilation of Hershberger assay data yielded 602 chemical/protocol/study combinations for 216 unique chemicals (Figure 1; Supplemental Table 2). Some chemical/study combinations were immediately excluded from analyses because they did not meet the defined parameters of the search (e.g. review articles, the test chemical dose could not be determined, studies were conducted in a different species or following a different test guideline). After a full-text review of the study design and data extraction, additional chemical/study combinations were subject to a second round of exclusion based on the criteria defined in Table 1. After the second round of exclusions, 371 chemical/study combinations for 134 chemicals remained for analyses (Figure 1). Independent studies were considered as a chemical tested in an experiment that included unique control groups.

3.1. Consistency of results for chemicals tested in >1 Hershberger study

Twenty-five chemicals were tested in two or more Hershberger studies that met our a priori criteria for including study results (Table 1) following the two step review process (see Supplemental Table 2, columns E, F and BN, BO for details). Of these 25 chemicals, 28% (7/25) had at least one disagreement (i.e. 1 positive and 1 negative study result; Tables 2 and 3, Supplemental Table 6). Many of the chemicals with multiple Hershberger studies were part of OECD validations and were used as androgenic, anti-androgenic, or negative reference chemicals in optimization of the guideline and the interlaboratory validation phases. Thirteen chemicals were tested in four or more Hershberger studies, including eight anti-androgens that varied in potency (p,p'-DDE, flutamide, finasteride, linuron, vinclozolin, procymidone, dibutyl phthalate, and prochloraz), three potent androgens (methyl

testosterone, testosterone propionate, and trenbolone), and two negative chemicals (4-nonylphenol and 2,4-dinitrophenol; Table 2). The range of no observed effect levels (NOEL) and lowest observed effect levels (LOEL) for all studies that met our inclusion criteria were reported for these 25 chemicals, along with the number of positive/negative study results for each protocol (Table 2). The NOELs and LOELs were extracted and reported for various subsets of Hershberger studies that: 1) adhered to the final OECD/US EPA guidelines (e.g. males were castrated between PND 42 and 53, allowed at 7 days post-surgical recovery, dosed for 10 consecutive days); 2) followed the stimulated weanling protocol (uncastrated males exposed to test chemical at PND 22 and co-treated with an androgen (methyl testosterone or testosterone propionate); and 3) fell under the category of “other” study protocols (e.g. different age of animal, different age a castration, use of intact non-weanling males, different in dosing duration), but otherwise met our criteria for inclusions (Table 2).

For 12 of 13 chemicals that were tested in four or more Hershberger studies, the range of LOEL concentrations overlapped with the range of NOEL concentrations (p,p'-DDE, flutamide, finasteride, linuron, 17-methyl testosterone, vinclozolin, procymidone, trenbolone, 4-nonylphenol, dibutyl phthalate (DBP), testosterone propionate, and prochloraz), and for linuron, procymidone, flutamide, and prochlorz, the ranges coincided entirely (Table 2). The sole exception was for the negative OECD reference chemical, 2,4-dinitrophenol which was negative in eight studies that met our criteria, and in seven of eight studies it was tested at a single concentration.

3.2. Comparison of Hershberger with other in vivo results

Sixty-five chemicals were run in at least one Hershberger study and in a male pubertal study or an in vivo study measuring androgen responsive endpoints reported in the same publication intended to support results of the Hershberger assay (Table 3; Supplemental Table 2). With the exception of methyl testosterone, which was androgenic in vivo, all chemicals with a Hershberger study and another in vivo study measuring androgen responsive endpoints were either negative or anti-androgenic. Of these, 30% (20/65) chemicals were negative in both the Hershberger assay(s) and the other in vivo assay, 32% (21/65) chemicals were positive (anti-androgenic) in both in vivo study designs, and 35% (23/65) chemicals had variable results such that calls could not be reliable made (Table 3; Supplemental Table 6). In addition, 43% (28/65) chemicals had some disagreement between the Hershberger and other in vivo assay results. As described previously, results from chemicals tested in more than one Hershberger study sometime disagreed (e.g. linuron, DBP, 4-nonylphenol, cyfluthrin, permethrin) and were also noted in Table 3. Of the 28 chemicals with disagreement in results, eight were anti-androgenic in 1 Hershberger study and negative in other in vivo assays, and 20 were negative in 1 Hershberger study and anti-androgenic in other in vivo assays (Table 3, Supplemental Table 6).

3.3. Chemicals with reproducible in vivo results in multiple studies that detect (anti)androgenic effects

Seventy-three unique chemicals were identified with either more than one Hershberger study (25 chemicals) or a Hershberger study and another in vivo study with androgen-responsive endpoints (65 chemicals, including 17 with 2 Hershberger study; Figure 1). From these 73

chemicals, 49 chemicals satisfied our criteria for consistent outcomes in vivo (positive effects in 2 in vivo studies, or no effects in 2 in vivo studies and no positive effect in any in vivo study; Table 3). The one exception to this logic was applied in the case of the OECD negative reference chemical 4-nonylphenol, which was negative in eight of nine validation trials (Table 2), and we considered to be overall negative for effects on the androgen pathway in vivo. The final list of chemicals with reproducible in vivo effects in multiple assays encompassed five androgenic chemicals, 23 anti-androgenic chemicals, and 21 negative chemicals (Figure 1; shaded green, red, and gray respectively in Table 3) and includes all OECD reference chemicals. In addition, 33% (24/73) chemicals had inconsistent results between in vivo tests and no consistent in vivo androgen pathway effect could be determined from these data (shaded yellow in Table 3, Supplemental Table 6), and 40% (29/73) chemicals had level of disagreement between in vivo study outcomes (Table 3, Supplemental Table 6).

3.4. Contribution of 5 α -reductase

The Hershberger assay should also detect 5 α -reductase inhibitors, though only one chemical with this mode of action was tested in the assay validation, there are no widely available assays that measure 5 α -reductase enzyme inhibition, and few 5 α -reductase inhibitors have been identified to date. To help identify chemicals that may have anti-androgenic effects in the Hershberger assay through this mechanism, a literature search was performed. A rank-order of citation frequency for 165 compounds correctly identified the most cited compounds as pharmaceuticals designed to directly inhibit 5 α -reductase for clinical applications, thereby supporting the suitability of the search key word selection. After cross-referencing the chemicals identified in 5 α -reductase literature search with compounds in the Hershberger database, only 12 chemicals were found in common (Supplemental Table 5). The pharmaceutical finasteride, an irreversible inhibitor of type II and III 5 α -reductase isozymes, was identified in 497 publications and was the most frequently cited chemical identified in the literature search (Supplemental Table 4). Three references were identified for progesterone, two references for both 17 α -ethinylestradiol and 17 α -estradiol, and one reference for the remaining eight Hershberger database chemicals (estrone, 17 β -estradiol, linuron, prochloraz, p,p'-DDE, fenarimol, flutamide, and 17-methyltestosterone; Supplemental Table 4). Based on the literature search, there is a gap in robust in vitro evidence to evaluate 5 α -reductase inhibition as a key mode of action of anti-androgenic chemicals.

4. Discussion

The Hershberger assay has been used as the standard for detecting in vivo (anti)androgenic effects for more than half a century and is part of the US EPA's battery of assays to screen for effects on the androgen pathway. In order to characterize uncertainty in the Hershberger assay and identify chemicals with reproducible in vivo effects on the androgen pathway, we developed a database of results from available scientific literature. We defined a priori criteria (Table 1) with the aim to include results of guideline Hershberger studies, as well as positive results from study designs determined to be less sensitive than the final test guideline, assuming that if a positive response was observed in a less sensitive model, it would likely be confirmed by a more sensitive model. As mentioned previously, animals

with intact HPG axes (e.g. stimulated weanlings and un-castrated adult males) may be less susceptible to AR-mediated effect due to possible compensatory mechanisms, and alternatively, may be susceptible to chemicals acting through other modes of action.

During the development and optimization of the test guideline, consideration was given to different protocols including age of animals, use of castrated or intact males, duration of post-surgical recovery, duration of exposure, and selection and dose of androgen co-administered in the anti-androgenic mode of the assay. While our a priori criteria for inclusion/exclusion were intentionally inclusive of positive results from various experimental designs, our analyses indicate a similar range of responses (e.g. NOELs, LOELs, positive/negative responses) among stimulated weanlings and castrated animals dosed for variable amounts of time (Table 2; Supplemental Table 2). This finding is consistent with other reported responses to anti-androgens among weanlings, peri-pubertal castrates, and adult castrates; though peri-pubertal males were more sensitive to weak anti-androgens (Ashby and Lefevre 2000a, OECD 2008a, OECD 2009c). Similar responses to exogenous androgens and anti-androgens were noted among castrated males with variable post-surgical recovery (Ashby and Lefevre 2000b), and among rats castrated at three, six, and 10 weeks of age (OECD 2008a; Yamada et al. 2001). Comparisons of results from the three models indicated peri-pubertal castrated rats were maximally sensitive to chemicals, because both AR and steroidogenic enzymes expression at this developmental stage result in high sensitivity to androgens and low AST weight, and thus a greater dynamic range and less individual variability that minimizes false negatives (OECD 2009b).

4.1 Variability in the HB

The primary objective of this investigation was to compile existing Hershberger data and evaluate the overall reproducibility of the assay. For the 25 chemicals with two or more Hershberger studies that met our criteria for evaluation, 28% (7/25) had disagreement between study results (Table 3). In some cases, this figure may underestimate the variability of chemical results because 13 of these 25 chemicals were reference chemicals in OECD validation studies, and were often tested at the same doses, using the same chemical stock, and strictly following guideline protocols. Despite being a much more complicated assay, the disagreements for chemicals tested in multiple Hershberger studies was about the same of that observed for the rat uterotrophic assay where 26% of chemicals tested in two or more studies had conflicting results, even when the study protocol adhered closely to the guideline method (Kleinstreuer et al. 2016).

As in the case of the variability in the uterotrophic results, differences in rat strain, study design and dosing likely contribute to the variability in results (Kleinstreuer et al. 2016). Comparison of Hershberger assays results for chemicals tested in more than one rat strain suggested differences in sensitivity among strains (OECD 2008a, Yamasaki et al. 2001a, You et al. 1998). For example, Sprague Dawley and Wistar rats were more sensitive to the AR antagonist, flutamide, than Fischer 344 rats (Yamasaki et al. 2001a). While strain differences may affect response to chemicals, 98% of the 371 studies included in our evaluations were tested in Sprague-Dawley or Wistar rats (286 and 78 studies, respectively),

though genetic differences in supplier and in-house rats of a single strain may contribute further to uncharacterized variability in responses (Yao et al. 2012).

Some the studies included in this investigation were part of the OECD method optimization and as such, there were differences in the doses of chemicals that were used as controls (e.g. flutamide, TP), co-administered androgen (e.g. TP), or reference chemicals (e.g. p,p'-DDE, linuron, nonylphenol). Variability in responses due to the dose is expected, though once the test method was optimized, well-characterized chemicals were often tested at the same doses or a single dose.

The most frequently tested chemical was the pesticide decomposition product, p,p'-DDE, used as a weak positive anti-androgen reference chemical in protocol optimization and method validation. As a result, the same p,p'-DDE chemical stock and dose levels were often used among the labs participating in the validation trials. Interestingly, the LOEL for p,p'-DDE ranged from 5 to 200 mg/kg bw/d and included concentrations below the low end of the range of reported NOEL values (10 to 50 mg/kg bw/d; Table 2; Supplemental Table 2). In some cases, this could be attributed to differences in the dose of androgens co-administered in the anti-androgenic mode of the assay or in the specific protocols, though these alone do not account for the variability in the response, and there is still overlap between the LOEL and NOEL when only results from guideline studies are considered (Table 2). For the 13 chemicals tested in four or more Hershberger studies, there is overlap between the LOEL and NOELs for all except the consistently negative reference chemical, 2,4-dinitrophenol; however, this is somewhat misleading because this chemical was always tested at a single concentration (Table 2).

Route of exposure is also a recognized source of variation, and while the Hershberger test guideline was designed to include oral or subcutaneous routes of exposure, an overwhelmingly proportion of chemicals were administered orally (95%; 354/371 studies). Only three chemicals (trenbolone, 2,4,4'-trihydroxy benzophenone, and testosterone propionate) were tested via both oral and subcutaneous administration. Trenbolone was administered orally in 10 of 11 Hershberger studies and by subcutaneous injection in one study. Though this was the only "non-guideline" study, this was due to castration occurring one day (PND 41) before the guideline recommendation (Table 2, Supplemental Table 2). Trenbolone resulted in significant increases in AST weights in all 11 tests, though the concentration at which an effect was observed was much higher when the chemical was administered subcutaneously (LOEL = 200 mg/kg bw/d, LOEC = 1.5–10 mg/kg bw/d when administered via oral gavage; Table 2, Supplemental Table 2). 2,4,4'-trihydroxy benzophenone was administered once orally and once subcutaneously. Oral administration of 2,4,4'-trihydroxy benzophenone resulted in no effect at the highest dose tested (600 mg/kg bw/d) compared to a significant effect observed at 300 mg/kg bw/d following subcutaneous dosing (Table 2, Supplemental Table 2). Testosterone propionate was administered via subcutaneous injection in five Hershberger assays on once by oral administration. The oral dosing study treated males castrated at PND 22, however, the LOEL was 0.1 mg/kg bw/d and within the range of LOELs observed in studies administering testosterone propionate via subcutaneous injection (Table 2, Supplemental Table 2). While route of test chemical exposure contributes to variability in study results, in this analysis,

dose route is a potential cause of inconsistent responses for only one (2,4,4'-trihydroxy benzophenone) of the seven chemicals tested in multiple Hershberger studies with variable in results.

For the six other chemicals with disparate results, a variety of factors may have contributed to disagreement among Hershberger responses. Cyfluthrin was tested in three Hershberger studies all conducted in Sprague-Dawley rats, though in two of the three, the males were castrated earlier (PND 28 versus PND 42–53) and dosed for a shorter duration (7 days versus 10 days) than recommended in the final guideline study (Table 1, Supplemental Table 2). Significant anti-androgenic effects were observed in the two studies of PND 28 castrates dosed for 7 days, but not in PND 42 castrated males dosed for 10 days (Supplemental Table 2). In the negative study, cyfluthrin was also co-administered with a lower dose of TP (0.4 mg/kg bw/d compared to 0.5 mg/kg bw/d in the positive studies), and the NOEL (20 mg/kg bw/d) was higher than the LOEL for one of the positive studies (18 mg/kg bw/d). An explanation for the differences in these results is not obvious, though it should be noted that cyfluthrin was also negative in a male pubertal rat study (Table 3; Supplemental Table 2). Dibutyl phthalate, administered orally in all studies, was anti-androgenic in six of seven Hershberger studies, all conducted in Sprague-Dawley males castrated on PND 42. The single negative study followed the guideline protocol, as did several positive studies. Curiously, three of DBP studies co-administered higher doses of TP (1–2 mg/kg bw/d) than in the negative study (0.4 mg/kg bw/d); though this same TP dose was also associated with detection of significant anti-androgenic effects in three studies (Supplemental Table 2). Dibutyl phthalate also had significant (anti-androgenic) effects in other in vivo assays (Table 3; Supplemental Table 2). Linuron, administered orally, was anti-androgenic in 17 of 18 Hershberger studies. The 17 positive studies include 10 studies following experimental designs that adhere to the guideline protocol and seven studies using the intact weanling model (Table 2; Supplemental Table 2). In six positive studies, 1.0 mg/kg bw/d TP was co-administered with linuron, versus 0.4 mg/kg bw/d TP in the negative study. The single negative study closely adhered to the guideline Hershberger protocol, however, the highest dose tested (100 mg/kg bw/d) was the upper end of the LOEL observed in all positive studies (10–100 mg/kg bw/d). While it is tempting to attribute the single negative result to not testing a high enough dose of test chemical, these results were part of the OECD validation and the same chemical, dose ranges, and protocols were used in multiple studies (Owens et al. 2007). Linuron has significant anti-androgenic effects in other in vivo protocols (Table 3; Supplemental Table 2). Permethrin was positive in two and negative in one Hershberger studies (Table 2; Table 3). In all three studies, permethrin was administered orally to castrated, Sprague-Dawley rats (Supplemental Table 2). Castration occurred at an earlier age in the two positive studies (PND 28 and 35, respectively). The negative results was from a study most closely following the guideline protocol, and surprisingly, the NOEL was 120 mg/kg bw/d, while the LOEL was 10 and 50 mg/kg bw/d, in the two positive studies (Supplemental Table 2). Co-administration of TP was 0.4 or 0.5 mg/kg bw/d in all three (and the 0.5 dose was associated with a positive result), and there is no clear explanation for the discrepancy between results of these studies. Permethrin significantly decrease serum testosterone levels but did not affect timing of pubertal (Supplemental Table 2). Bifenthrin was negative in one Hershberger study that closely adhered to the guideline

protocol and positive in one study using animals castrated on PND 28. In both cases, Sprague-Dawley rats were dosed orally. In this case, the highest dose tested in the negative study was 10 mg/kg bw/d while the LOEL was 13.5 mg/kg bw/d, and therefore, it is possible that the negative study did not dose high enough. 4-nonylphenol was considered a negative reference chemical in multiple phases of the OECD Hershberger assay validation and was negative in eight of nine Hershberger assay results (Tables 2 and 3). Because it was used in assay validation, the experimental designs of all nine studies closely followed the guideline protocol (and each other) and the single positive result is difficult to explain.

Closer consideration of chemicals with inconsistencies in Hershberger study results provide insight into possible explanations for only two of the seven chemicals (2,4,4'-trihydroxy benzophenone due to difference in dosing route; bifenthrin due to low test chemical concentration). There is a possibility that differences in genetically undefined rat strains, chemical stocks, difference in excision and measuring techniques, or a combination of all factors may contribute to variability in the results of multiple studies for a single chemical. Reproducibility of the Hershberger assay is very similar to other estimates of reproducibility for the uterotrophic (Kleinstreuer et al. 2016) and rodent local lymph node assay (Hoffmann et al. 2018).

In several ways, this analysis may overestimate the reproducibility of Hershberger study results. For example, our criteria for including negative study results were very stringent and if chemicals were tested at only one or two (in the anti-androgenic mode of the assay) doses, or only four AST weights were examined and one weight was significantly different from controls, these studies were excluded during the second round of data review. If this level of stringency on negative test results is relaxed to include all guideline protocols including those tested at one or two doses and measuring four or more AST weights, the number of chemicals with inconsistencies between study results increases to 53% (16/30; Supplemental Table 2).

4.2. HB comparison with in other in vivo studies

Of the 65 chemicals tested in at least one Hershberger assay and another in vivo assay with androgen responsive endpoints, 65% (42/65) had significant effects in both types of assays (22) or lack of effects in both assays (20). Although the majority of results indicated a consistent effect, there was at least one conflicting Hershberger assay result for five of the chemicals with reproducible results (Table 3). For 43% (28/65) chemicals tested in a Hershberger assay and another in vivo study, including 23 EDSP List 1/Tier 1 chemicals, some disagreement between results was observed (US EPA 2015b; Table 3). One might expect chemicals that act as potent AR agonists or antagonists (e.g. pharmaceuticals) to also have effects in assays that use animals with intact HPG axes, but effects of less potent chemicals detected in the Hershberger may be diminished by compensatory mechanisms. Conversely, chemicals with no effects in the Hershberger assay and positive effects in other in vivo studies may be acting through mechanisms other than the AR that are not well detected by the Hershberger assay. For example, atrazine, 4-nonylphenol, and tebuconazole were negative in the Hershberger and positive in HPG-intact animals, and act through well characterized modes of action on central feedback regulation, estrogen receptors, and

steroidogenesis, respectively. Differing results of chemicals tested using different animal models can be explained by biology; however, this may introduce some difficulties for interpreting overall weight of evidence effects on the androgen pathway and role of the Hershberger assay data in that interpretation.

4.3. Chemicals with reproducible in vivo results

From the initial identification of over 200 chemicals tested in Hershberger studies, 49 chemicals (5 androgenic, 23 anti-androgenic, and 21 negative chemicals) showed reproducible androgen pathway effects that could be confirmed in more than one in vivo study (Table 3). Despite the extensive effort required to identify studies, review experimental designs, and extract experimental data, such undertakings are needed to identify chemicals with reproducible in vivo effects. Robust reference chemicals can be used to interrogate alternatives to animal tests (Kleinstreuer et al. in prep). The hope is that these undertakings can be helpful for model building and may be expanded in the future.

Similar to the results from the uterotrophic database (Kleinstreuer et al. 2017), the degree of variability in the Hershberger assay was relatively high. Though only 25 chemicals were tested in multiple Hershberger studies, 13 of the chemicals were used in test guideline development and validation. Despite controlling for several potential sources of uncertainty in validation studies, results from chemicals used in the validation appeared to have a range in responses similar to the other 12 chemicals tested in more than one study. Results between Hershberger studies and other in vivo study designs were also variable. Although it is difficult to compare the Hershberger studies with other in vivo study designs due to different sensitivities to specific modes of action, the level of disagreement observed within the Hershberger assays with strict limitations on experimental design (i.e. Hershberger versus Hershberger; 28%) was unexpectedly high compared to the disagreements between Hershberger studies and other in vivo experimental designs (43%).

4.4. Interpretation of the HB

OECD guidance for evaluating chemicals for endocrine disruption suggests the Hershberger can be used as part of an integrated testing strategy following a positive result in an in vitro AR assay to confirm a positive response in vivo, or to evaluate an AR mode of action for a response observed in a higher tiered test in whole animals (OECD 2012). The Hershberger assay is part of the EDSP Tier 1 screening battery and interpreted along with the Tier 1 in vitro AR binding assay, male pubertal rat assay, fish short-term reproduction assay (FSTRA), and other available data to screen for potential effects of environmental chemicals on the androgen pathway. Though the EDSP Tier 1 assays each provide specific information (e.g. AR binding assay indicates mechanism of action, the FSTRA is an aquatic exposure of animals with an intact neuroendocrine axis and functional steroidogenesis), there is deliberate redundancy between endpoints from different assays (US EPA 2011).

Ankley and Gray (2013) examined the consistency of responses between the EDSP Tier 1 guideline Hershberger and male pubertal rat assays for six chemicals with well-characterized endocrine effects. Five chemicals that were positive in the Hershberger study were also positive in the male pubertal study, and one chemical, methoxychlor, was negative in the

Hershberger assay and positive in at least one male pubertal study (Ankley and Gray 2013). The authors suggested the pubertal male assays is important for confirming observations in the Hershberger assays, as well as identifying chemicals that may disrupt endocrine function by different mechanisms (Ankley and Gray 2013).

Here, we expand on the Ankley and Gray analysis, examining 65 chemicals tested in at least one Hershberger study and a male pubertal study or other in vivo rat study used to confirm results of the Hershberger assay in the same publication. A total of 43% (28/65) of chemicals showed some disagreement (Table 3). Of these, 26% (17/65) of chemicals had no effect in the Hershberger assay(s) but had significant effects in male rats with intact HPG axes (e.g. delay timing of puberty; altered AST weights or peripheral testosterone levels). Animals with a functional neuroendocrine axis are sensitive to a variety of other modes of action such as altered steroidogenesis, perturbation of other endocrine pathways or alteration of central feedback regulation that effect results and make comparison with results of the neuroendocrine-interrupted Hershberger model difficult. When there was an indication of anti-androgenic effects of chemicals from Hershberger results and an absence of effects in the male pubertal assay for eight EDSP List 1/Tier 1 chemicals (or data accepted in lieu of that assay), five chemicals were determined by US EPA to have no potential androgen pathway activity (Table 3). The three remaining chemicals that were determined to have potential androgen pathway effects also had positive results in another in vivo assay (e.g. FSTRA or other rat study). Given these results, it is difficult to determine what is gained from running the Hershberger assay as part of the EDSP Tier 1 battery.

The challenges associated with interpreting the results from the Hershberger assay may stem from the fact that it is not simply an androgen bioassay. In the Hershberger validation studies, some chemicals were hypothesized to act as “anti-androgens” by inducing hepatic enzymes and increasing TP clearance, resulting in decreased AST weights among TP + test chemical treated animals relative to controls (TP only) animals, rather than acting directly through the AR (US EPA 2015b, Freyberger and Schladt 2009, Fryeberger et al. 2007). In contrast to the three-day uterotrophic assay, the longer Hershberger assay provides sufficient time for a test chemical to increase xenobiotic enzyme expression and activity and may substantially contribute to endpoint responses. In one such example, Marty et al. (2014) noted three EDSP List1/Tier 1 chemicals that were negative for AR binding in Tier 1 and ToxCast assays, had AST weight changes that were not consistent with 5 α -reductase inhibition (e.g. no differential reduction in prostate versus other tissues), and caused a 2.5 fold induction of hepatic glucuronosyltransferases (UGTs), suggesting the anti-androgenic response was not through a direct AR mode of action. In this case, serum testosterone levels were decreased by 29% relative to control animals, though this change was not statistically significant (Marty et al. 2014). Benfluralin significantly reduced AST weights at concentrations that also enhanced hepatic testosterone clearance, suggesting the observed anti-androgenic effect was due to enhanced androgen clearance (US EPA 2015b).

Our analyses indicated 22% (30/134) of chemicals that were included after our two-round vetting process caused significant increases in liver weight in at least one Hershberger study and nine chemicals (7%) caused increases in liver weight 25% (Supplemental Table 2). While we considered excluding chemicals that caused increases in liver weights from our

analyses, several of these (e.g. methyl testosterone, finasteride, flutamide, linuron, procymidone, vinclozolin) were reference chemicals used to optimize and validate the assay. The most commonly tested chemical, the weak AR reference antagonist p,p'-DDE, caused significant increases in liver weights in 73% (41/56) of the studies in which it was tested. Chemicals that affect the liver of animals tested in the Hershberger study would likely affect the livers of rats used in pubertal or other in vivo test methods. However, for animals with intact HPG axes, a variety of compensatory mechanisms may help to dispel anti-androgenic effects due to increased endogenous testosterone. Further, because those experimental designs do not co-administer test chemicals with a synthetic androgen, increased hepatic clearance may not result in an "anti-androgenic" decrease in AST weight.

The anti-androgenic portion of the Hershberger assay is a key distinction from the uterotrophic assay and introduces a substantial complication. Selection and doses of the co-administered reference androgen probably differentially affect ASTs and the magnitude of the response, factors that may have significant implications for observing the effects of test compounds. In comparing experimental designs during OECD guideline development and validation, results indicated differential responsiveness of ASTs to the same reference androgen among rats of different ages, and different sensitivities to different reference androgens among rats of the same age (Ashby and Lefevre 2000b, OECD 2008a, OECD 2009a; Supplemental Table 2). For all four chemicals with multiple positive (anti-androgenic) test results and at least one negative test result, the negative result was in a study with TP co-administered in the anti-androgenic phase of the test at 0.4 TP mg/kg bw/d (which adheres to the guideline recommendations). In these instances, at least one positive result was reported with a higher co-administration of TP (ranging from 0.5 to 2.0 mg/kg bw/d; see Section 4.1, Supplemental Table 2). This is somewhat counter intuitive; as one might hypothesize the higher dose of androgen could overwhelm anti-androgenic effects of the test chemical, and from a limited evaluation, benchmark dose for decreased AST weights were lower for lower dose (0.2 versus 0.4 mg/kg bw/d) of co-administered TP (Owens et al. 2007). Rather, at these higher doses, TP may contribute to xenobiotic enzyme induction and rapid clearance, and therefore, are detected as "anti-androgenic".

In addition to functional livers, rats in the Hershberger assay are capable of responding to chemicals that inhibit 5 α -reductase activity. The AST weights in the Hershberger assay have differential responses to testosterone and dihydrotestosterone (DHT). For example, ventral prostate weight depends on local 5 α -reductase activity to convert testosterone to DHT, whereas SV weights are less dependent on DHT, and the LABC has essentially no ability to convert testosterone to DHT (Tyl et al. 2007; OECD 2008b). These differences can be illustrated in the response of ASTs to trenbolone, a potent androgen that will bind to and activate the AR but is not a substrate for 5 α -reductase. Trenbolone treatment reportedly increased GP, LABC, and CG weights, but not VP or SV weights in previous studies (Freyberger et al. 2005). In the evaluation of Hershberger studies measuring effects of trenbolone, all five AST weights were significantly increased in nine of 11 studies evaluated herein, calling into question if changes in specific AST weights can be interpreted as a potential indicator of chemical mechanism (Supplemental Table 2). Similarly, calculations of AST-specific benchmark dose from labs included in OECD assay validation studies

indicated as much as an order of magnitude difference in doses for the five ASTs measured (Table 5, Owens et al. 2007).

The literature search for chemicals with Hershberger study results that might affect 5 α -reductase activity only yielded information on 12 chemicals (Supplemental Table 7). One reason for the limited number of chemicals identified may be due to the difficulty distinguishing in vivo anti-androgenic effects due to 5 α -reductase inhibition (and decreased endogenous DHT production) from anti-androgenic effects mediated through the AR, and the coincident absence of a reliable, widely available in vitro assay to elucidate the chemical mechanism of action. One study tested chemicals with known in vivo anti-androgenic effects on DHT synthesis using two independent human enzyme assays (Lo et al. 2007). Results of the enzyme assays indicated that p,p'-DDE, linuron, fenarimol, and flutamide inhibited 5 α -reductase activity at high concentrations (IC₅₀ 24 μ M) in human prostate homogenates and had no activity in lymph node carcinoma of prostate (LNCaP) cells (Lo et al. 2007). Methyltestosterone and prochloraz were more potent inhibitors of DHT production (IC₅₀ 1.9 and 12.4 μ M, respectively) in prostate homogenates and considerably higher in LNCaP cells (Lo et al. 2007). Other studies used a variety of disparate methods which are difficult to compare. Nukui (1997) reported weak inhibition of in vitro conversion of testosterone to DHT (K_i = 15 μ M) in human epididymal microsomes made from patients treated with 17 α -ethinyl estradiol. Differences were noted by epididymal region, and the author hypothesized that the reduction in 5 α -reductase activity was due to effects on protein expression rather than substrate inhibition (Nukui 1997). 17 α -estradiol was noted to reduce 5 α -reductase mRNA in rats following exposure to a high fat diet (Cai et al. 2011) and reverse alopecia in women undergoing treatment (Blume-Peytavi et al. 2007), but enzyme activities were not reported in these cases. Preliminary data suggested estrone, 17 β -estradiol, and progesterone were progressively more potent inhibitors of 5 α -reductase (K_i = 15.5, 5.1, and 0.11 μ M, respectively) but V_{max} and K_m indicated endogenous concentrations likely have minimal effects in vivo (Kreig et al. 1985). This is supported by doses of 100 nM of 17 α -estradiol and 17 β -estradiol inhibiting DHT production in human hair follicles by 20% and 60% respectively (Niiyama et al. 2001), and gram-level doses of progesterone being required to elicit in vivo effects on AST weights (de Larminat and Blaquier et al. 1979). Though these data are limited by the availability of a reliable assay, result of our literature search suggest that, other than finasteride, only a few chemicals may have effects on 5 α -reductase, but at concentrations much greater than their well characterized estrogen receptor, androgen receptor, progesterone receptor or cytochrome P450 enzyme inhibition modes of action.

Lastly, the Hershberger data considered in this manuscript were typically generated following the OECD/US EPA Hershberger guideline or as part of the initial optimization and validation of the guideline. In most cases, these studies included two or fewer dose groups for androgenic effects and three or fewer dose groups for anti-androgenic effects (plus appropriate control animals). The NOEL and LOEL levels can be determined from these studies but the current guideline is not adequate for reliable calculation of benchmark dose from a few treatment groups that may cover a wide range of concentrations. The imprecision of NOEL/LOEL concentrations based on relatively small sample sizes and wide separation between adjacent doses may substantially contribute to uncertainty in the Hershberger results.

Conclusions

The intention of this project was to develop a database of Hershberger studies and other in vivo studies with AR-related endpoints, characterize results of chemicals tested in multiple Hershberger assays, and to identify chemicals with reproducible in vivo results for the purposes of serving as reference chemicals that can be used for developing/evaluating alternative methods. The results of our analysis indicated a high degree of variability in results of Hershberger studies. Among chemicals tested in multiple Hershberger assays and meeting our criteria for evaluation, 28% of chemicals had disagreements between the study results (i.e. 1 positive and 1 negative study) and in most cases, these differences cannot be explained by study protocol or dose level. In addition, we also found a lack of consistency between Hershberger assay conclusions and other in vivo assay measuring androgen-responsive endpoints. Most chemicals that have androgen pathway effects act as anti-androgens and “anti-androgenic” Hershberger effects may be difficult to recapitulate in other in vivo studies. The anti-androgenic mode of the Hershberger assay is tested in rodents that are co-administered an androgen (e.g. TP) to increase sensitivity of the AST response. Many chemicals may induce hepatic xenobiotic metabolizing enzymes, resulting in increased clearance of the TP suggesting an anti-androgenic effect. Other in vivo rodent study designs do not include co-administration of the test chemical with an androgen and thus, the effect is not corroborated in some cases. Based on the limited reproducibility of the Hershberger assay, it may be difficult to draw conclusions for chemicals. In fact, chemicals screened in US EPA’s EDSP Tier 1 battery were only considered to have potential androgen pathway effects if a positive Hershberger assay result was corroborated by an additional positive in vivo study result.

The added value of Hershberger data in an overall weight of evidence evaluation of a chemical’s potential effect on the androgen pathway is not clear. Negative data do not necessarily rule out AR mediated effects. The co-administration of TP, alone, is enough to induce hepatic enzymes and may increase metabolism (and thus limit the effect) of the test chemical. Similarly, if the test chemical is a weak anti-androgen it may not be sufficiently potent to overwhelm effects of TP, a pharmaceutical androgen. Organ weight changes are not sensitive indicators of AR interactions, and therefore, lack of a positive Hershberger result does not completely rule-out AR mediated test chemical effects. A positive effect may be further difficult to interpret because the guideline protocol uses HPG-interrupted animals, and thus, results may not be recapitulated in other in vivo assays. If AR screening is the goal of the Hershberger assay, this could be more easily addressed by an AR in vitro assay, albeit one that could account for hepatic metabolism.

Finally, we identified 49 chemicals with reproducible androgen pathway effects confirmed in more than one in vivo study. The hope is that these chemicals and database may be used to develop alternative methods of androgen pathway screening.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

The authors gratefully acknowledge the assistance of the ICCVAM reference chemical working group and Dr. LE Gray for assistance in identifying criteria for high quality animal studies, Drs. S Matten and K Hamernik of the EPA Office of Science Coordination and Policy for input on the Hershberger literature curation, and Drs. M DeVito (NTP), LE Gray (US EPA), and R. Hines (US EPA) for technical review of this manuscript.

Disclaimer. The authors declare no competing financial interests. This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors. This manuscript and the view expressed herein are those of the authors and do not necessarily reflect the views or policies of the US EPA, NIH, or OECD.

LITERATURE CITED

- Ankley GT, Gray LE. 2013 Cross-species conservation of endocrine pathways: a critical analysis of tier 1 fish and rat screening assays with 12 model chemicals. *Environ Tox Chem.* 32:1084–1087.
- Ashby J, and Lefevre PA 2000a The peripubertal male rat assay as an alternative to the Hershberger castrated male rat assay for the detection of anti-androgens, oestrogens and metabolic modulators. *J. Appl. Toxicol* 20, 35–47. [PubMed: 10641015]
- Ashby J, and Lefevre PA 2000b Preliminary evaluation of the major protocol variables for the Hershberger castrated male rat assay for the detection of androgens, antiandrogens, and metabolic modulators. *Regul. Toxicol. Pharmacol* 31, 92–105.
- Ashby J, Owens W, Deghenghi R, and Odum J. 2000a Concept evaluation: an assay for receptor mediated and biochemical antiestrogens using pubertal rats. *Reg. Tox. Pharm* 35:393–397.
- Ashby J, Owens W, and Lefevre PA. 2002 Concept evaluation: androgen-stimulated immature intact male rats as an assay for antiandrogens. *Reg. Tox. Pharm.* 35:280–285.
- Ashby J, Lefevre PA, Tinwell H, Odum J, Owens W. 2004 Testosterone-stimulated weanlings as an alternative to castrated male rats in the Hershberger anti-androgen assay. *Reg Toxicol Pharm.* 39:229–238.
- Blume-Peytavi U, Kunte C, Krisp A, Garcia Bartels N, Ellwanger U, Hoffmann R. 2007 Comparison of the efficacy and safety of topical minoxidil and topical alfatradiol in the treatment of androgenic alopecia in women. *J. Dtsch Dermatol Ges.* 5: 391–395. [PubMed: 17451383]
- Browne P, Judson RS, Casey WM, Kleinstreuer NC, Thomas RS. 2015 Screening Chemicals for Estrogen Receptor Bioactivity Using a Computational Model. *Environ Sci Technol* 49(14): 8804–8814 [PubMed: 26066997]
- Bruns RF, Watson IA. 2012 Rules for identifying potentially reactive or promiscuous compounds. *J Med Chem* 55, 9763–9772. [PubMed: 23061697]
- Cai LQ, Cai J, Wu W, Zhu YS. 2011 17 α -estradiol and genistein inhibit high fat diet induced prostate gene expression and prostate growth in the rat. *J. Urol* 186: 1489–1496. [PubMed: 21855913]
- De Larminat ME, Blaquier JA. 1979 Effect of in vivo administration of 5 alpha reductase inhibitors on epididymal function. *Acta Physiol Lat Am.* 29:1–6. [PubMed: 263219]
- Ezendam J, Braakhuis HM, Vandbriel RJ. 2016 State of the art in non-animal approaches for skin sensitization testing: from individual test methods towards testing strategies. *Arch Toxicol* 2016 9 14 [Epub ahead of print]
- Freyberger A, Hartmann E, Krotlinger F. 2005 Evaluation of the rodent Hershberger bioassay using three reference (anti)androgens. *Arch Hig Rada Toksikol.* 56:131–139. [PubMed: 15968828]
- Freyberger A, Ellinger-Ziegelbauer H, Krotlinger F. 2007 Evaluation of the rodent Hershberger bioassay: testing coded chemicals and supplementary molecular-biological and biochemical investigations. *Toxicology.* 239:77–88. [PubMed: 17688994]
- Freyberger A, Schladt L. 2009 Evaluation of the rodent Hershberger bioassay on intact juvenile males- testing of coded chemicals and supplementary biochemical investigations. *Toxicology.* 262:114–120. [PubMed: 19467291]
- Hershberger LG, Shipley EG, Meyer RK. 1953 Myotrophic activity of 19-nortestosterone and other steroids determined by modified levator ani muscle methods. *Proc Soc Exp Biol Med.* 83:175–180. [PubMed: 13064212]

- Hoffmann S, Kleinstreuer NC, Alépée N, Allen D, Api AM, Ashikaga T, Clouet E, Cluzel M, Desprez B, Gellatly N, Göbel C, Kern PS, Klaric M, Kühnl J, Lalko JF, Martinuzzi-Teissier S, Mewes K, Miyazawa M, Parakhia R, van Vliet E, Zang Q, Petersohn D. 2018 Non-animal methods to predict skin sensitization (I): the Cosmetics Europe database. *Crit Rev Toxicol*. 48:344–358. [PubMed: 29474128]
- Kleinstreuer NC, Ceger P, Watt ED, Martin M, Houck K, Browne P, Thomas RS, Casey WM, Dix DJ, Allen D, Sakamuru S, Xia M, Huang R, Judson R. 2017 Development and validation of a computational model for androgen receptor activity. *Chem Res Toxicol* 30:946–964. [PubMed: 27933809]
- Kleinstreuer NC, Ceger PC, Allen DG, Strickland J, Chang X, Hamm JT, et al. 2016 A Curated Database of Rodent Uterotrophic Bioactivity. *Environ Health Persp* 124(5): 556–562.
- Lo S, King I, Allera A, Klingmuller D. 2007 Effects of various pesticides on human 5 α -reductase activity in prostate and LNCaP cells. *Toxicology In Vitro* 21: 502–508 [PubMed: 17218080]
- Marty MS, O'Connor JC. 2014 Key learnings for the Endocrine Disruptor Screening Program (EDSP) Tier 1 uterotrophic and Hershberger assays. *Birth Defects Res B Dev Reprod Toxicol*. 101:63–79. [PubMed: 24515841]
- NRC. 2007 Toxicity testing in the 21st century: A vision and a strategy. National Academies Press: Washington, DC; Vol. 25, p 136–138.
- Niiyama S, Happle R, Hoffmann R. 2001 Influence of estrogens on the androgen metabolism in different subunits of human hair follicles. *Eur J Dermatology*. 11:195–198.
- Nukui F 1997 Effects of chlormadinone acetate and ethinylestradiol treatment on epididymal 5 α -reductase activities in patients with prostate cancer. *Endocrine Journal* 44: 127–132. [PubMed: 9152625]
- OECD. 2007 Draft Report of the OECD Validation of the Rat Hershberger Bioassay: Phase 3: Multiple Laboratories. Surgical Castrate Model Protocol. ENV/JM/TG/EDTA(2007)2; 16-2-2007 pp. 1–125.
- OECD 2008a Background Review Document on the Rodent Hershberger Bioassay. Number 90. Series on Testing and Assessment. ENV/JM/MONO(2008)17. [http://www.oecd.org/officialdocuments/publicdisplaydocumentpdf/?cote=env/jm/mono\(2008\)17&doclanguage=en](http://www.oecd.org/officialdocuments/publicdisplaydocumentpdf/?cote=env/jm/mono(2008)17&doclanguage=en).
- OECD. 2008b Report on the OECD Validation of the Hershberger Bioassay: Phase 2: Testing of Androgen Agonists, Androgen Antagonists, and a 5 α -Reductase Inhibitor in Dose Response Studies by Multiple Laboratories. OECD Series on Testing and Assessment, Number 86.
- OECD. 2008c Detailed Review Paper on the Use of Metabolising Systems for In Vitro Testing of Endocrine Disruptors. Series on Testing and Assessment. Number 97. [http://www.oecd.org/officialdocuments/publicdisplaydocumentpdf/?cote=env/jm/mono\(2008\)24&doclanguage=en](http://www.oecd.org/officialdocuments/publicdisplaydocumentpdf/?cote=env/jm/mono(2008)24&doclanguage=en)
- OECD 2009a OECD Guidelines for Testing of Chemicals, Section 4. Test No. 441: Hershberger bioassay in rats. A short-term screening assay for (anti)androgenic properties. 10.1787/9789264076334-en
- OECD 2009b Report on the validation of the Hershberger bioassay (weanling model). Number 108 Series on Testing and Assessment. ENV/JM/MONO(2009)13. [http://www.oecd.org/officialdocuments/publicdisplaydocumentpdf/?cote=env/jm/mono\(2009\)13&doclanguage=en](http://www.oecd.org/officialdocuments/publicdisplaydocumentpdf/?cote=env/jm/mono(2009)13&doclanguage=en)
- OECD. 2009c Guidance Document on the Weanling Hershberger Bioassay in Rats: A Short-term Screening Assay for (Anti)androgenic properties. Number 115 Series on Testing and Assessment. ENV/JM/MONO(2009)41. <http://www.oecd.org/officialdocuments/publicdisplaydocumentpdf/?cote=env/jm/mono%282009%2941&doclanguage=en>
- OECD 2012 Guidance Document on Standardised Test Guidelines for Evaluating Chemicals for Endocrine Disruption. Series on Testing and Assessment, No. 150. <http://www.oecd.org/officialdocuments/publicdisplaydocumentpdf/?cote=env/jm/mono%282012%2922&doclanguage=en>.
- Owens W, Gray LE, Seiger E, Walker M, Yamasaki K, Ashby J, Jacob E. 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 Perspective* 115: 671–678.

- US EPA. 1998 Endocrine Disruptor Screening and Testing Advisory Committee Final Report (1998). Available at: <http://www.epa.gov/endocrine-disruption/endocrine-disruptor-screening-and-testing-advisory-committee-edstac-final>.
- US EPA. 2007 Integrated Summary Report for Validation of a test method for assessment of pubertal development and thyroid function in juvenile male rats as a potential screen in the Endocrine Disruptor Screening Program Tier-1 Battery. Available at: <https://www.regulations.gov/document?D=EPA-HQ-OPP-2008-0012-0017>.
- US EPA. 2009a Endocrine Disruptor Screening Program Test Guidelines. OPPTS. 890.1400: Hershberger Bioassay. Available at: <https://www.regulations.gov/document?D=EPA-HQ-OPPT-2009-0576-0008>.
- US EPA. 2009b Endocrine Disruptor Screening Program Test Guidelines. OPPTS. 890.1400: Uterotropic Assay. Available at: <https://www.regulations.gov/document?D=EPA-HQ-OPPT-2009-0576-0012>.
- US EPA. 2009c U.S. EPA. Series 890-Endocrine Disruptor Screening Test OPPTS 890.1500: pubertal development and thyroid function in intact juvenile/peripubertal male rats. Available at: <http://www.regulations.gov/#!documentDetail;D=EPA-HQ-OPPT-2009-0576-0010>.
- US EPA. 2009d EPA's approach for considering Other Scientifically Relevant Information (OSRI) under the Endocrine Disruptor Screening Program. Available at: <http://www.regulations.gov/#!documentDetail;D=EPA-HQ-OPPT-2007-1080-0032>.
- US EPA. 2011 Weight-of-Evidence: Evaluating results of EDSP Tier 1 screening to identify the need for Tier 2 testing, U. S. Environmental Protection Agency (2011). Available online at: <http://www.regulations.gov/#!documentDetail;D=EPA-HQ-OPPT-2010-0877-0021> [Accessed 03/08/16].
- US EPA. 2012 Universe of Chemicals and General Validation Principles, U.S. Environmental Protection Agency, Endocrine Disruptor Screening Program, Washington DC Available online at: http://www.epa.gov/sites/production/files/2015-07/documents/edsp_chemical_universe_and_general_validations_white_paper_11_12.pdf
- US EPA. 2015a Use of High Throughput Assays and Computational Tools: Endocrine Disruptor Screening Program; Notice of Availability and Opportunity for Comment, 80 Fed. Reg. 118 (6 19, 2015). Office of Chemical Safety and Pollution Prevention, U.S. Environmental Protection Agency, Washington DC Available online at: <https://www.federalregister.gov/articles/2015/06/19/2015-15182/use-of-high-throughput-assays-and-computational-tools-endocrine-disruptor-screening-program-notice>
- US EPA. 2015b Endocrine Disruptor Screening Program Tier 1 Assessments. Office of Chemical Safety and Pollution Prevention, U.S. Environmental Protection Agency, Washington DC Available online at: <https://www.epa.gov/ingredients-used-pesticide-products/endocrine-disruptor-screening-program-tier-1-assessments>
- Yao CQ, Prokopec SD, Watson JD, Pang R, P'ng C, Chong LC, Harding NJ, Pohjanvirta R, Okey AB, Boutros PC. 2012 Inter-strain hereogeneity in rat hepatic transcriptomic responses to 2,3,7,8-tetrachlorodibenzo-p-dioxin (TCDD). *Toxicol Appl Phramacol.* 260: 135–45.
- Yamada T, Sunami O, Kunimatsu T, Kamita Y, Okuno Y, Seki T, Nakatsuka I, Matsuo M (2001). "Dissection and weighing of accessory sex glands after formalin fixation, and a 5-day assay using young mature rats are reliable and feasible in the Hershberger assay", *Toxicology*, 162: 103–19. [PubMed: 11337110]
- Yamasaki K, Sawaki M, Takatsuki M. 2001 Strain sensitivity differences in the Hershberger assay. *Reprod Toxicol.* 15:437–440. [PubMed: 11489600]
- Yamashita S 2004 Localization of estrogen and androgen receptors in male reproductive tissues of mice and rats. *Anat Rec A Discov Mole Cell Evol Biol.* 279: 768–778.
- You L, Casanova M, Archibeque-Engle S, Sar M, Fan LQ, Heck HA. 1998 Impaired male sexual development in perinatal Sprague-Dawley and Long-Evens hooded rats exposed in utero and lactationally to p,p'-DDE. *Toxicol Sci.* 45:162–173. [PubMed: 9848123]

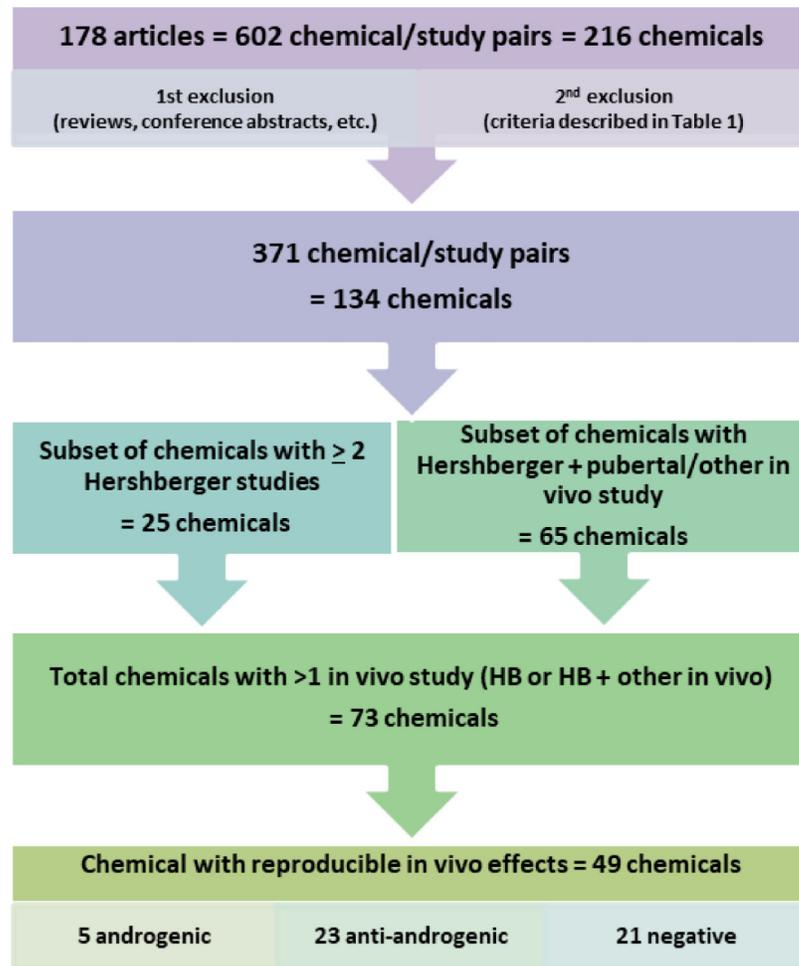


Figure 1. Workflow and subsets of chemicals identified from the original systematic literature review of Hershberger and other in vivo studies with androgen-responsive endpoints.

Table 1.

Methodological descriptors for the Guideline Hershberger assay and criteria considered here for positive in vivo androgenic and anti-androgenic results and negative results with no significant in vivo effect on >1 accessory sex tissue weight.

Parameter	Guideline	Positive results	Negative results
Animal model	male rats castrated PND 42 to 53	castrated males, intact stimulated weanlings, males castrated at earlier time point	same as guideline
Castration/recovery	7 days post-surgical recovery before dosing	same as guideline	same as guideline
Sample size	4 animals/group for positive results; 6 animals/group for negative results	3 animals per group	same as guideline
Dosing route	oral gavage or SC	oral gavage, SC, IM, IP	same as guideline
Androgenic doses	2 test chemical dose levels	1 dose level	same as guideline
Androgenic controls	positive control: TP 0.2–0.4 mg/kg bw/d; negative control: vehicle	positive control TP, TC, T, MT; negative control: vehicle	same as guideline
Anti-androgenic doses	3 test chemical dose levels + coadministration with TP 0.2–0.4 mg/kg bw/d	1 dose level + coadministration with TP, TC, T, MT = 100	same as guideline
Anti-androgenic controls	positive control: TP 0.2–0.4 mg/kg bw/d + 3 mg/kg bw/d flutamide; negative control: TP 0.2–0.4 mg/kg bw/d	positive control flutamide 3 mg/kg bw/d, p,p'-DDE; negative control: TP, TC, T, MT	same as guideline
Dosing duration	10 days beginning PND 49 to 60	5 days	same as guideline
AST measured	VP, SV, L/ABC, CG, GP	2 of 5 measured	4 of 5 measured
Criterion for positive/negative determination	Significant effects in 2 AST compared to appropriate control	same as guideline	lack of significant effects in 4 of 5 required AST
Necropsy	18–36 hrs after last dose	same as guideline	same as guideline

PND = postnatal day, SC = subcutaneous, IM = intramuscular, IP=intraperitoneal, TP= testosterone propionate, TC = testosterone propionate, MT= methyl testosterone, VP = ventral prostate, SV=seminal vesicles, L/ABC=levator ani bulbocavernosus, CG=coagulating gland, GP= glans penis, AST=accessory sex tissue

Table 2.

Ranges in no observed effect level (NOEL) and lowest observed effect level (LOEL) for 25 chemicals with 2 Hershberger studies included after the second exclusion filter was applied (see text and Table 1 for details).

Chemical/study design (n=number of studies)	NOEL (mg/kg bw/d)	LOEL (mg/kg bw/d)	# Pos/Neg results	OECD reference classification
p,p'-DDE (n=56)	10–50	5–200		Anti-androgenic
Guideline* (n= 29)	10–50	16–200	29/0	
Stimulated weanling (n=9)	16–50	5–160	9/0	
Other models (n=18)	10–30	30–200	18/0	
flutamide (n=32)	0.15–0.6	0.1–100		Anti-androgenic
Guideline (n=18)	-	1–100	18/0	(AR antagonist)
Stimulated weanling (n=2)	-	0.1–3	2/0	
Other (n=12)	0.15–0.6	0.6–100	12/0	
finasteride (n=28)	0.002–0.2	0.008–25		Anti-androgenic
Guideline (n=19)	0.002–0.2	0.008–25	19/0	(5αR inhibitor)
Stimulated weanling (n=7)	0.008	0.04–25	7/0	
Other (n=2)	-	25	2/0	
linuron (n=18)	10–100	10–100		Anti-androgenic
Guideline (n=11)	10–100	10–100	10/1	
Stimulated weanling (n=7)	10–30	100	7/0	
Other	-	-		
17-methyl testosterone (n=20)	0.1–20	0.5–100		Androgenic
Guideline (n=12)	0.1–10	0.5–50	12/0	
Stimulated weanling	-	-		
Other (n=8)	0.5–20	5–100	8/0	
vinclozolin (n=20)	3–30	10–100		Anti-androgenic
Guideline (n=10)	3–30	10–100	10/0	
Stimulated weanling (n=2)	-	30	2/0	
Other (n=8)	3–10	10–100	8/0	
procymidone (n=14)	3–30	3–100		Anti-androgenic
Guideline (n=11)	3–10	3–100	11/0	
Stimulated weanling (n=2)	3–30	10–100	2/0	
Other (n=1)	3	10	1/0	
trenbolone (n=11)	1.5–100	8–200		Androgenic
Guideline (n=10)	1.5–10	8–50	10/0	
Stimulated weanling	-	-		
Other (n=1)	100	200	1/0	

Chemical/study design (n=number of studies)	NOEL (mg/kg bw/d)	LOEL (mg/kg bw/d)	# Pos/Neg results	OECD reference classification
4-nonylphenol (n=9)	160–200	160		Negative
Guideline (n=9)*	160–200	160	1/8	
Stimulated weanling	-	-		
Other	-	-		
dibutyl phthalate (n=7)	500	500–1000		NA
Guideline (n=6)	500	500–1000	5/1	
Stimulated weanling	-	-		
Other (n=1)	500	1000	1/0	
2,4-dinitrophenol (n=8)	10	-		Negative
Guideline (n=8)†	10	-	0/8	
Stimulated weanling	-	-		
Other	-	-		
testosterone propionate (n=6)	0.02–0.1	0.0125–0.25		Androgenic
Guideline (n=4)	0.02–0.1	0.0125–0.2	4/0	
Stimulated weanling	-	-		
Other (n=2)	0.06	0.1–0.25	2/0	
prochloraz (n=4)	62.5	50–250		NA
Guideline (n=1)	62.5	125	1/0	
Stimulated weanling	-	-		
Other (n=3)		50–250	3/0	
testosterone (n=3)	0.1	0.05–0.2	-	NA
Guideline	-	-	-	
Stimulated weanling	-	-	-	
Other (n=3)	0.1	0.05–0.2	3/0	
fenarimol (n=3)		200		NA
Guideline	-	-	-	
Stimulated weanling	-	-	-	
Other (n=3)	-	200	3/0	
cyfluthrin (n=3)	6–20	18–50		
Guideline (n=1)	20	-	0/1	NA
Stimulated weanling	-	-		
Other (n=2)	6	18–50	2/0	
permethrin (n=3)	120	10–50		
Guideline (n=1)	120	-	0/1	NA

Chemical/study design (n=number of studies)	NOEL (mg/kg bw/d)	LOEL (mg/kg bw/d)	# Pos/Neg results	OECD reference classification
Stimulated weanling	-	-	-	
Other (n=2)	-	10–50	2/0	
3-(dibutylamino)phenol (n=3)				NA
Guideline (n=3)	400	-	0/3	
Stimulated weanling	-	-		
Other	-	-		
methyl-1-testosterone (n=2)	0.03	0.3–1		NA
Guideline (n=1)	0.03	0.3	1/0	
Stimulated weanling	-	-		
Other (n=1)	-	1	1/0	
beta-cyfluthrin (n=2)	12	36–50		NA
Guideline	-	-		
Stimulated weanling	-	-		
Other (n=2)	12	36–50	2/0	
bis(2-thylhexyl)phthalate (DEHP) (n=2)	20–100	100–200		
Guideline (n=1)	20	100	1/0	NA
Stimulated weanling	-	-		
Other (n=1)	100	200	1/0	
2,4,4'-trihydroxy benzophenone (n=2)	100–600	300		
Guideline (n=1)	600	-	0/1	NA
Stimulated weanling	-	-		
Other (n=1)	100	300	1/0	
bifenthrin (n=2)				
Guideline (n=1)	10	-	0/1	NA
Stimulated weanling	-	-		
Other (n=1)	-	13.5	1/0	
3-amino-1,2,4-triazole				
Guideline (n=2)	1000	-	0/2	NA
Stimulated weanling	-	-		
Other	-	-		
4,4'-butylidenebis(2-tert-butyl-5-methylphenol) (n=2)				
Guideline (n=2)	1000	-	0/2	NA
Stimulated weanling	-	-		
Other	-	-		

*“Guideline” studies closely followed the OECD 441/EPA 890.1350 guideline (i.e. males were castrated between postnatal day (PND) 42 and 53, allowed 7 d post-surgery, dosed for 10 d). “Stimulated weanling” studies were conducted on uncastrated males, stimulated with an androgen and treated with test chemical. “Other models” deviated from the guideline study (e.g. castrated before or after PND 42–53, dosed <10 d, dosing began <7 d after castration), but adhered to criteria for evaluating positive and negative study results defined in Table 1.

†4-nonylphenol and 2,4-dinitrophenol were included as negative chemicals and were typically only tested at a single dose in the androgenic or anti-androgenic study designs. Thus, though all other criteria are met, they do not meet the requirements for minimum test chemical doses with the exception of one study for 4-nonylphenol that tested three doses (in which the test chemical was inactive).

EPA Author Manuscript

EPA Author Manuscript

EPA Author Manuscript

Table 3.

Summary of findings for 73 unique chemicals with in vivo results for chemicals with >1 Hershberger study (25) or at least 1 Hershberger study and a male pubertal or other in vivo study (65). Shaded areas indicate chemicals with reproducible in vivo effects (green=androgenic, red=anti-androgenic, grey=negative; yellow = inconsistent effects). Also indicated are chemicals with disagreements between multiple Hershberger studies or Hershberger and other in vivo studies. In the case where there were disagreements between multiple Hershberger tests, the chemical was described by the predominant activity (e.g. dibutyl phthalate was anti-androgenic in 6/7 studies and permethrin was anti-androgenic in 2/3 studies).

Chemical Name	CAS	Androgenic in HB	Anti-androgenic in HB	NE in HB	Total HB studies	disagreement in >1 HB	Tested in other in vivo study	Androgenic in HB and other	Anti-androgenic in HB and other	Anti-androgenic in HB; NE in other	NE in HB, Anti-androgenic in other
17-methyl testosterone	58-18-4	20			20			x			
methyl-1-testosterone	65-04-3	2			2						
trenbolone	10161-33-8	11			11						
testosterone	58-22-0	3			3						
testosterone propionate	57-85-2	6			6						
benfluralin	1861-40-1		1		1				x		
beta-cyfluthrin *	68359-37-5		2		2						
bis(2-ethylhexyl) phthalate (DEHP)	117-81-7		2		2				x		
cyfluthrin	68359-37-5		2	1	3	x				x	
DE-71	32534-81-9		1		1				x		
dibutyl phthalate (DBP)	84-74-2		6	1	7	x			x		(x)
ethoprop	13194-48-4		1		1				x		
fenarimol	60168-88-9		3		3						
fenbutatin oxide	13356-08-6		1		1				x		
finasteride	98319-26-7		28		28				x		
flutamide	13311-84-7		32		32				x		
iprodione	36734-19-7		1		1				x		

Chemical Name	CAS	Androgenic in HB	Anti-androgenic in HB	NE in HB	Total HB studies	disagreement in >1 HB	Tested in other in vivo study	Androgenic in HB and other	Anti-androgenic in HB and other	Anti-androgenic in HB; NE in other	NE in HB, Anti-androgenic in other
linuron	330-55-2		17	1	18	x			x		(x)
metolachlor	51218-45-2		1		1				x		
noflurazon	27314-13-2		1		1				x		
p,p'-DDE	72-55-9		56		56				x		
permethrin	52645-53-1		2	1	3	x				x	
prochloraz	67747-09-5		4		4				x		
procymidone	32809-16-8		15		15				x		
pronamide	23950-58-5		1		1				x		
propargite	2312-35-8		1		1				x		
trifluralin	1582-09-8		1		1				x		
vinclozolin	50471-44-8		20		20				x		
2,4,4'-trihydroxy benzophenone	1470-79-7		1	1	2	x					
acephate	30560-19-1		1		1					x	
atrazine	1912-24-9			1	1						x
bifenthrin	82657-04-3		1	1	2	x				x	
bisphenol A	80-05-7			1	1						x
captan	133-06-2			1	1						x
carbaryl	63-25-2			1	1						x
carbendazim	10605-21-7			1	1						x
chlorthal dimethyl (DCPA)	1861-32-1			1	1						x
cypermethrin	52315-07-8		1		1					x	
diazinon	333-41-5		1		1					x	
dichlobenil	1194-65-6		1		1					x	

Chemical Name	CAS	Androgenic in HB	Anti-androgenic in HB	NE in HB	Total HB studies	disagreement in >1 HB	Tested in other in vivo study	Androgenic in HB and other	Anti-androgenic in HB and other	Anti-androgenic in HB; NE in other	NE in HB, Anti-androgenic in other
imidacloprid	138261-41-3			1	1						x
isophorone	78-59-1			1	1						x
malathion	121-75-5			1	1						x
methoxychlor	72-43-5			1	1						x
myclobutanil	88671-89-0			1	1						x
o-phenylphenol	90-43-7			1	1						x
phosmet	732-11-6		1		1					x	
piperonyl butoxide	51-03-06			1	1						x
pyriproxifen	95737-68-1			1	1						x
simazine	122-34-9			1	1						x
tebuconazole	107534-9-63			1	1						x
triademifon	43121-43-3			1	1						x
3-(dibutylamino)phenol	43141-69-1			3	3						
3-amino-1,2,4-triazole	61-82-5			2	2						
2,4 -dinitrophenol	51-28-5			8	8						
4,4'-butylidenebis(2-tert-butyl-5-methylphenol)	85-60-9			2	2						
4-nonylphenol	25154-52-3		1	8	9	x			(x)		x
abamectin	71751-41-2			1	1						
acetone	67-64-1			1	1						
carbofuran	1563-66-2			1	1						
chlorothalonil	1897-45-6			1	1						
chlorpyrifos	2921-88-2			1	1						
esfenvalerate	66230-04-4			1	1						
flutolanil	66332-96-5			1	1						

Chemical Name	CAS	Androgenic in HB	Anti-androgenic in HB	NE in HB	Total HB studies	disagreement in >1 HB	Tested in other in vivo study	Androgenic in HB and other	Anti-androgenic in HB and other	Anti-androgenic in HB; NE in other	NE in HB, Anti-androgenic in other
folpet	133-07-3			1	1						
glyphosate	1071-83-6			1	1						
metalayl	57837-19-1			1	1						
methomyl	16752-77-5			1	1						
metribuzin	21087-64-9			1	1						
mgk-264	113-48-4			1	1						
oxamyl	23135-22-0			1	1						
PCNB	82-68-8			1	1						
tetrachlorvinfos	22248-79-9			1	1						

* beta-cyfluthrin is an optical isomer of cyfluthrin. The CASRN is that of cyfluthrin.