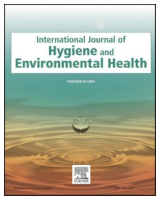
International Journal of Hygiene and Environmental Health 241 (2022) 113942

Contents lists available at ScienceDirect

International Journal of Hygiene and Environmental Health

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



Bisphenol A and declining semen quality: A systematic review to support the derivation of a reference dose for mixture risk assessments

Andreas Kortenkamp *, Olwenn Martin, Sibylle Ermler, Asma Baig, Martin Scholze

Brunel University London, Centre for Pollution Research and Policy, College of Health, Medicine and Life Sciences, Kingston Lane, Uxbridge, UB8 3PH, UK

ARTICLE INFO

Keywords:
Bisphenol A
Systematic review
Semen quality
Reference doses
Mixture risk assessment

ABSTRACT

To support a mixture risk assessment with a focus on male reproductive health, we conducted a systematic review of associations between bisphenol A (BPA) exposures and declines in semen quality, based on animal and epidemiological studies. Contrary to a widely held view that there is "conflicting" evidence of such associations, our review and confidence rating approach reveals that animal studies provide convincing evidence of declines of semen quality after gestational BPA exposures. Many of the reported negative findings can be attributed to deficiencies in study sensitivity, insufficient control of background contamination and probable confounding through hormonal interference due to the use of soy-containing diets. We did not evaluate animal studies of adult BPA exposures. Divergent findings in "medium to high" and "medium" confidence epidemiological studies can be explained in terms of differences in exposure conditions. We attempted the estimation of a BPA reference dose based on animal studies. Due to variations in the no-observed adverse effect levels (NOAELs) in high confidence studies, possible reference doses ranged from 0.0001 to 0.0099 μg/kg/d. In choosing 0.003 μg/kg/d we struck a balance between caution suggested by studies at the lower end of the doses and the weight of evidence from studies with higher NOAELs. This weighting was motivated by the intended use of the value in a mixture risk assessment which meant arriving at a *reasonable estimate* of BPA exposures likely without effects on semen quality. We realise that our approach does not conform with the standards necessary for deriving tolerable daily intakes (TDIs) for single chemical exposures, which is not our interest here. BPA exposures currently experienced by European populations and beyond are in excess of 0.003 μg/kg/d and even fall in the range where some epidemiological studies observed effects on semen quality as a result of BPA exposures in adulthood.

1. Introduction

Bisphenol A (BPA) is a widely used industrial chemical that can disrupt several hormone systems and produce a variety of toxic effects. As a monomer in polycarbonate plastics and epoxy resins, it leaches into food and liquids. Polycarbonates are widely used as food contact materials, bottles and other containers. Epoxy resins make up the protective lining inside food cans and the coatings of water pipes and tanks. BPA is also present in thermal paper used as cash receipts. While BPA use in baby bottles is now prohibited in the European Union, it is still permitted in food contact material and containers, with a migration limit of 50 ng/g food. As a result, BPA exposure is widespread and food items stored in cans (e.g. fish, tomatoes) and plastic bottles (e.g. milk) contribute significantly to the daily intake of the general population (Karrer et al., 2020).

The European Chemicals Agency (ECHA) has included BPA in the list of substances of very high concern on the basis of endocrine disrupting properties (ECHA, 2018). Evidence has been mounting that BPA interferes with the signalling processes important for healthy male reproductive development. As is common with many endocrine disruptors, BPA affects multiple endpoints that constitute a syndrome of effects related to poor male reproductive health. It can antagonise the binding of testosterone to the androgen receptor (AR) (Ermler et al., 2011). After exposure of rats during gestation it produces changes in the anogenital distance (AGD) of male offspring (Christiansen et al., 2014) and declines in semen quality (Hass et al., 2016), both indicators of diminished androgen action in fetal life. Several epidemiological studies report associations between BPA exposures in adult life and declines in several parameters of semen quality (Adoamnei et al., 2018; Ji et al., 2018; Lassen et al., 2014).

The precise mechanisms by which BPA affects semen quality are not resolved, but several possibilities can be envisaged. Cell-autonomous activation of the AR in Sertoli cells is an absolute requirement for

* Corresponding author.
E-mail address: andreas.kortenkamp@brunel.ac.uk (A. Kortenkamp).

https://doi.org/10.1016/j.ijheh.2022.113942
Received 13 November 2021; Received in revised form 11 January 2022; Accepted 2 February 2022
Available online 12 February 2022

Abbreviations

AF	Assessment factor
AGD	Anogenital distance
AR	Androgen receptor
AUC	Area under the curve
BPA	Bisphenol A
COSTER	Conduct of Systematic Reviews in Toxicology and Environmental Health Research
ECHA	European Chemicals Agency
EFSA	European Food Safety Authority
ELISA	Enzyme-linked immunosorbent assay
GD	Gestational day
HED	Human equivalent dose
HEDF	Human equivalent dose factor
HI	Hazard Index
LOAEL	Lowest observed adverse effect level
NOAEL	No observed adverse effect level
NTP	National Toxicology Program
OHAT	Office of Health and Translation
PECO	Population, Exposure, Control, Outcome
PND	Postnatal day
TDI	Tolerable daily intake
t-TDI	temporary tolerable daily intake
WHO	World Health Organisation

Sertoli cells to be able to support spermatogenesis and to allow germ cells to complete meiosis (De Gendt et al., 2004). The ability of BPA to disrupt germ cell meiosis by producing abnormal proportions of stages VII – VIII of the spermatogenic cycle (Shi et al., 2018) could therefore be attributed to its AR-antagonist properties. However, there are other effects, including disruption of the epigenetic programming necessary for spermatogenesis, as evidenced by expression changes in DNA methyl transferases (Shi et al., 2019) and increased oxidative stress in testicular tissues (Ullah et al., 2019).

Experimental studies have shown that BPA can act in combination with other AR antagonists *in vitro* (Orton et al., 2014). In multi-component mixture studies of gestational exposures in the rat, BPA acted in concert with anti-androgens to produce retained nipples in male offspring (Axelstad et al., 2014) and declines in semen quality (Axelstad et al., 2018).

Numerous other chemicals can also affect male reproductive health, including phthalates, parabens, analgesics, polychlorinated dioxins, polychlorinated biphenyls, polybrominated biphenyl ethers and certain azole pesticides (Kortenkamp, 2020). Exposures to these substances are widespread (Apel et al., 2020; Moos et al., 2017; Bauer et al., 2021; Martin et al., 2017). This calls for systematic investigations of the impact of simultaneous exposure to multiple chemicals on male reproductive health.

One widely used mixture risk assessment method is the Hazard Index (HI) (Teuschler and Hertzberg, 1995). It employs risk quotients of exposure and health-based guidance values or reference doses that are familiar from single chemical risk assessment. By summing up the risk quotients of all chemicals included in the mixture risk assessment it examines fold-exceedances of "acceptable" mixture exposures relative to an index value of 1. To achieve consistency of the mixture risk assessment, these risk quotients must be built with reference doses for similar toxicity endpoints. Utilisation of reference doses for different toxicities, e.g. carcinogenicity for one mixture component and lung toxicity for another, must be avoided as such mixing of toxicities increases the uncertainty of the assessment. Thus, the search for the most sensitive toxicity endpoint, the so-called critical toxicity, which is required for deriving health-based guidance values in single chemical risk assessments, is not the sole criterion in mixture risk assessments. It must be complemented by estimating doses associated with a common adverse outcome.

Apart from disrupting male sexual differentiation, BPA affects a multitude of other processes, with adverse outcomes. In their recent Draft Scientific Opinion on BPA, the Panel on Food Contact Materials, Enzymes and Processing Aids of the European Food Safety Authority (EFSA) identified the immune system as the most sensitive target of BPA exposure. BPA also produces metabolic effects, developmental neurotoxicity and adverse effects on female reproductive organs. To protect the immune system from BPA exposures, the EFSA Panel derived a new tolerable daily intake (TDI) of 0.04 ng/kg body weight/day (EFSA, 2021), considerably lower than the previous temporary TDI of 4 μg/kg body weight/day (EFSA, 2015). However, a value derived for immunotoxicity cannot be relied on for a mixture risk assessment for male reproductive health. It is therefore necessary to derive a BPA reference dose specifically for reproductive effects.

In view of the widespread declines in semen quality in Western countries (Levine et al., 2017), and with the intention of interpreting these unfavourable trends in the framework of a mixture risk assessment, we selected semen quality as the basis for deriving a BPA reference dose.

We conducted a systematic review of the epidemiological literature and of the body of evidence from animal experimental studies to address two separate but related questions: what is the strength of evidence of associations between BPA exposure and declines in semen quality? What is a BPA reference dose for semen quality declines that can be used in a mixture risk assessment of male reproductive health, specifically with a focus on semen quality? We placed particular emphasis on gestational BPA exposures because germ stem cell populations are established in fetal and neonatal life, and only after this period spermatogenesis can begin. Disruption of these processes can have life-long, irreversible effects. In the mouse, this period is from gestational day (GD) 7 to postnatal day (PND) 8, in the rat from GD 9 to PND 10 (de Rooij and Vergouwen 1991; Olaso and Habert 2000). For obvious reasons, it is difficult to establish accurately such periods in humans, but the equivalent window is presumed to be in the first trimester of pregnancy (Sharpe, 2020).

2. Materials and methods

2.1. Literature search and screening

Literature search and screening, study evaluation, data extraction and evidence synthesis methods are set out in detail in the systematic review protocol developed following the COSTER recommendations (Whaley et al., 2020; Martin et al., 2021; and Supplementary Material S1). Briefly, epidemiological studies and experimental studies with BPA describing declines in semen quality were identified by conducting literature searches in PubMed, Web of Science, Scopus until July 2020, updated in August 2021. Citation searches of key papers were also conducted. PECO statements and literature search algorithms are available in Supplementary Material S1.

We incorporated all experimental studies with laboratory animals that analysed total sperm count, sperm concentration, motility, morphology or vitality as outcome measures, but did not consider DNA damage or aneuploidy. Studies with non-mammalian species were excluded, as were studies where BPA was administered to adult animals, outside the period when germ cell stem populations are established between GD 7 to PND 10. We also excluded studies where BPA was injected (sub-cutaneously or intra-peritoneally) as these routes bypass

liver metabolism and can lead to inflated BPA tissue concentrations.

We included epidemiological studies among adult men (between 18 and 40 years of age) that reported semen quality parameters (total sperm count, sperm concentration, motility, morphology or vitality). Case-control studies, cohort studies and cross-sectional studies were considered, but case reports and reviews were excluded. Only studies that had assessed BPA exposure as urinary concentrations were eligible. Measurements in plasma, serum or cord blood were excluded, due to the absence of kinetic models that allow estimation of daily intakes based on the concentrations measured in these fluids. Studies that measured BPA concentrations by using ELISA were also deemed unreliable and were not considered. Studies reporting associations between BPA and DNA damage in sperm, or aneuploidy were also excluded, as these effects are not related to disruptions of male reproductive health by hormonal factors.

The literature review process was coordinated and managed using the freely available CADIMA tool (https://www.cadima.info/index.php/area/evidenceSynthesisDatabase). Title/abstract, full text screening and data extraction was performed by at least two reviewers.

2.2. Study evaluation

The internal validity (risk of bias) of individual studies was assessed using separate criteria and considerations for human epidemiological and for animal studies. Our main concerns were risk of bias (understood as factors that affect the magnitude or direction of effect) and insensitivity (factors that limit the ability of a study to detect an effect that is in fact present).

To appraise the internal validity of experimental studies with mammalian laboratory animals, we used the internal validity appraisal protocol (risk of bias assessment) for BPA as detailed in (EFSA, 2017) and (EFSA, 2019). EFSA developed this protocol following the NTP OHAT Risk of Bias Tool (described in the NTP OHAT 2019 Handbook for conducting a literature-based health assessment, p 33 (NTP, 2015). We used key elements similar to those defined by EFSA (2019) for appraising BPA studies and analysed exposure characterisation (purity of test compound, consistent administration, and absence of contamination of the test compound), outcome assessment (blinding of assessors) and power of detecting effects (sufficient numbers of animals per dose group).

To assess specific quality issues related to studies of BPA and semen quality, we introduced three further key elements in our appraisal. One of these concerns the control of BPA contamination by using polycarbonate-free caging. BPA can leach from polycarbonate caging (Howdeshell et al., 2003) and may thus obscure the effects of experimentally administered BPA. Second, the use of phytoestrogen-free chow is important as phytoestrogen-containing chow may introduce hormonal disturbances which mask the effects of BPA on semen quality (Ruhlen et al., 2011). The third additional key element concerns the inclusion of a positive control with established detrimental effects on semen quality (often ethinylestradiol, estradiol or diethylstilbestrol). This is necessary to demonstrate the proficiency of the investigators to detect changes in semen quality and shows that the experimental system is sufficiently sensitive. The complete list of appraisal elements can be found in the published protocol (https://doi.org/10.5281/zenodo.5083147, Supplementary Material S1) and in Table 1.

Each element was scored using the NTP OHAT categories ++ *Definitely low risk*, + *Probably low risk*, ~ *Probably high risk* and ~~ *Definitely high risk*. Key elements were assessed first, and a study that failed a key element was not evaluated further. We adopted the system in EFSA (2019) and rated each study in terms of three Tiers, with *Tier 1* signifying high confidence where all three EFSA key elements and all our three additional key elements are scored + or ++ and no more than 1 question not addressing these key elements is scored ~ or ~~ (see EFSA, 2019; Table 2, p 8). *Tier 2* signifies medium confidence and denotes all combinations not covered in Tier 1 or 3. Studies were placed in *Tier 3* (low confidence) when any one of the three EFSA key elements and the additional key elements was scored ~ or ~~ or when more than 50% of the questions not addressing these key elements were scored ~ or ~~. The risk of bias assessment protocol is shown in the published protocol, together with instructions how to rate each element of the protocol in terms of the risk categories.

We examined epidemiological studies of associations between BPA and semen quality using the procedures detailed by Radke et al. (2018), with evaluations of exposure measurement, outcome measurement, participant selection, confounding and analysis. By applying the criteria detailed in Radke et al. (2018) and listed in the published protocol (https://doi.org/10.5281/zenodo.5083147) we judged the quality of each study regarding its utility for hazard identification by reaching a consensus in each evaluation domain with the categories *Good*, *Adequate*, *Poor*, or *Critically Deficient*. The ratings for each evaluation domain were combined to obtain an overall study confidence rating of *High*, *Medium*, *Low*, or *Uninformative* (Table 2).

2.3. Data synthesis

We provided a narrative synthesis to summarise the characteristics and findings of the eligible studies, in terms of BPA exposure ranges not associated with declines in human studies, or in terms of NOAELs or LOAELs in animal studies. In these summaries we only considered human epidemiological studies rated as high or medium confidence, and experimental animal studies rated as high confidence (*Tier 1*). To enable quantitative comparisons between bisphenol A exposures in human studies and experimental studies with animals, we converted urinary bisphenol A levels into daily intakes for humans by employing the toxicokinetic model detailed in Koch et al. (2012).

2.4. Evidence synthesis

We first assessed whether the evidence linking BPA with declines in semen quality, from both human and animal studies, is sufficiently robust to support hazard identification. To address this question, we employed methods for weighing evidence from two lines of evidence, human and animal studies, following the principles described in EFSA guidance (EFSA 2017). The evidence was synthesised by considering aspects of an association that may suggest causation, according to the Bradford Hill criteria, based on EFSA's adaptation: consistency, exposure–response relationship, strength of association, temporality, biological plausibility, and coherence.

We synthesised evidence from animal studies and human epidemiological studies separately to derive a reference dose for mixture risk assessment. For animal experiments we utilised the framework in Radke et al. (2018) modified by the approach detailed in EFSA (EFSA, 2019), as follows: The evidence is categorised as *Robust* when there are sets of studies with a *Tier 1* confidence rating with consistent findings of adverse effects on semen quality across multiple laboratories and species. Any evidence that cannot be reasonably explained by the respective study design or differences in animal model is from a set of experiments of lower confidence (*Tier 2* or *Tier 3*). The category *Moderate* is assigned when a set of evidence does not reach the degree of certainty required for *Robust*, but which includes at least one *Tier 1* confidence study and information strengthening the likelihood of a causal association. The results are largely consistent, but notable uncertainties remain. *Slight* describes a scenario in which there is a suggestion of a possible effect on semen quality, but the evidence is conflicting or weak, with only *low* confidence experiments available. *Indeterminate* is used when no animal studies are available or where the evidence is highly inconsistent and primarily of *low* confidence. *Compelling evidence of no effect* is used when *high* confidence experiments demonstrate a lack of biologically significant effects across multiple species, both sexes, and a broad range of exposure levels.

We synthesised evidence from human studies by adopting the

Table 1
Evaluation of experimental animal studies and semen quality after BPA administration.

	Study description		Key appraisal elements						Study outcomes		Study evaluation	
Reference	Species	Outcome measures	Purity of chemical	Diet soy free?	Randomisation, concealment, blinding	Number of animals per group	Sensitivity of model, positive control	Background contamination with BPA?	Outcomes	Comments	Tier	Overall confidence
Cagen et al. 1999	Rat, Wistar	daily sperm production	>99%	no	not reported	25 to 28	DES, ineffectve	not reported	No effect		3	Low
Chatsantiprapa et al. 2016	Mouse, CD1	number of motile sperm	99%	not reported	not reported	12	none	not reported	Decrease in number of motile sperm	AGDi shortening at 50 ug/kg d but, not at 500 ug/kg d	1	High
Chiocarelli et al. 2020	Mouse, CD1	% viable sperm, % motile	99%	yes	yes	5	none	no	Decrease in live and motile sperm		1	High
Delclos et al. 2014	Rat, SD	sperm counts, motility	>99%	yes	yes	18 to 23	EE2, active	no	No effect		1	High
Dere et al. 2018	Rat, SD	Spermatid heads	>99%	yes	yes	10 to 20	none	no	No effect		3	Low
Ema et al. 2001	Rat	Number of sperm, % motile, % progressive motile, % abnormal, % tailless	>99%	not reported	yes	25	none	no	No effect	No effects on sperm counts or other sperm parameters; **decrease** in abnormal and tailless sperm at 20 ug/kg d, AGD changes	3	Low
Hass et al. 2016	Rat, Wistar	Number of sperm	>99.5%	yes	yes	17 to 21	none	no	Decrease in sperm number	Effect not seen at higher doses.	1	High
Howdeshell et al. 2008	Rat, LE	Number of sperm	>99%	no	yes	16 to 18	EE2, active	possible, due to polycarbonate cages	No effect		3	Low
Kendig et al. 2012	Mouse, CD1	Number of sperm, % motile	USEPA / NIEHS standard	yes	yes	8	EE2, ineffective	no	No effect	Sperm counts and motility **increased** at higher doses	3	Low
Kobayashi et al. 2010	Mouse, C57BL/6J	Number of sperm, % motile sperm	>99.6%	no	yes	12	none	not reported	Decrease in motile sperm	Sperm motility in F2 males was only impacted at the highest dose. Doses are approximated from ppm to ug/kg/d and were higher at some timepoints	3	Low
Kobayashi et al. 2012	Rat, SD	Number of sperm, % motile, % progressive motile	>99.6%	no	not reported	10	none	not reported	No effect on motility	Sperm counts not assessed or not given	3	Low
Meng et al. 2018	Mouse, C57BL/6	Number of sperm, % abnormal sperm	>99%	no	not reported	7	none	not reported	Decrease in sperm number, increase in malformed sperm		3	Low
Nagao et al. 2002	Mouse, C57BL/6N	Number of sperm	>99%	yes	not reported	25	none	not reported	No effect	The only effect observed was a decrease in the weight of the seminal vesicle at the highest dose	3	Low
Rahman et al. 2017	Mouse, CD1	Sperm conc, % motile	>99%	yes	not reported	3, in 3 independent experiments	none	no	Decrease in sperm conc		1	High
Salian et al. 2009	Rat, Hol	Number of sperm, % motile	>99%	yes	yes	24	DES, active	not reported	Decrease in sperm number and motility		1	High
Shi et al. 2018	Mouse, CD1	Sperm conc; % motile	>99%	not reported	yes	5	none	not reported	Decrease in sperm number		1	High
Shi et al. 2019	Mouse, CD1	Sperm conc, motility	>99%	not reported	yes	6	none	not reported	Decrease in sperm number and motility		1	High
Sporndly-Nees et al. 2018	Rat, F344	Morphologically abnormal sperm	>99%	yes	yes	8 to 12	none	no	No effect	Long lag between end of exposure (PND 21) and measurement of sperm morphology (12 months), sperm numbers not measured	3	Low
Tinwell et al. 2002	Rat, SD, Alderley	Number of sperm	>99%	no	Females assigned to dose groups based on body weight ranking	7 litters	EE2, active	possible, due to polycarbonate caging	Decrease in sperm number	Effect only seen in Alderley Park rats, not in SD rats	3	Low
Tyl et al. 2008	Mouse, CD1	Number of sperm, % motile, % abnormal	99.70%	no	not reported	10 to 28	E2, inactive	not reported	No effect	High levels of phytoestrogens measured in diet	3	Low
Tyl et al. 2002	Rat, SD	Number of sperm, % motile, % abnormal	99.5% pure	no	yes	10 to 30	none	no	Decrease in sperm number	Decrease on sperm counts at highest dose	3	Low
Ullah et al. 2019	Rat, SD	Number of sperm, % motile, % viable	>99%	yes	yes	8	none	no	Decrease in sperm number		1	High
Vilela et al. 2014	Mouse, Vesper	Number of sperm	>99%	yes	not reported	15 or 11 for two lowest doses, 8 for highest	DES, active	not reported	Decrease in sperm number and motility		1	High
Vom Saal et al. 1998	Mouse CF1	Number of sperm	not reported	no	not reported	5 to 7	none	no	Decrease in sperm number		3	Low
Yang et al. 2015	Rat, SD	Number of sperm, % motile	>99%	not reported	not reported	3	none	not reported	Decrease in sperm number and motility		1	High
Yoshino et al. 2002	Rat, F344	Number of sperm, % motile, morphology	99.90%	not reported	yes	5 to 10	none	not reported	No effect		3	Low

Table 2
Evaluation of epidemiology studies of associations of BPA with semen quality.

Reference	Study description			Semen quality outcomes					Study evaluation					
	Population	Exposure sampling	Outcome	N	C	Mot	Mor	Vit	Exposure	Outcome	Participant selection	Confounding	Analysis	Overall confidence
Adoamnei et al. 2018	Cross-sectional; Students, 18 - 23 y	spot urine; adjusted for creatinine	Number, concentration, motility, morphology	v	v	~	~	nd	P	G	G	G	G	M/H
Benson et al. 2021	Cross-sectional; young men from Danish National Birth Cohort	single spot urine sample, corrected for creatinine	Number, concentration, motility, morphology	~	~	~	~	nd	P	G	G	P	G	L
Caporossi et al 2020	Cross-sectional; male partners of sub-fertile couples	spot urine; adjusted for creatinine	Number, concentration, motility, morphology	~	~	~	~	nd	P	G	A	G	G	M
Chen et al. 2013	Case-control; infertile men, idiopathic	morning urine, adjusted for creatinine	Number, concentration	~	~	nd	nd		P	P	A	G	G	L
Den Hond et al. 2015	case-control; male patients of fertility clinics	spot urine, adjusted for creatinine	Number, concentration, motility, morphology	~	~	~	~	nd	P	G	A	P	A	L
Goldstone et al. 2015	Cohort; Men > 18 y	spot urine; adjusted for creatinine	Number, concentration, motility, morphology	~	~	~	~	nd	P	A	G	G	G	M
Pollard et al. 2019	Cohort; Men 18 - 40 y	first morning urine; 8 to 9 samples per subject; adjustment for creatinine	Number, concentration, morphology	~	~	nd	v	nd	A	A	G	P	A	M
Ji et al. 2018	Cross-sectional; men 18 - 55 y	single spot urine, adjusted for creatinine	Number, concentration, motility	~	v	v	nd	nd	P	G	G	G	G	M/H
Kim, Ko et al. 2019	Cross-sectional; male patients of fertility clinic	single spot urine, corrected for specific gravity	Concentration, motility	nd	~	~	nd	nd	P	P/CD	A	P	P	U
Knez et al. 2014	Cross-sectional; men fertility clinic, 34 y average	single spot morning urine, adjusted for creatinine	Number, concentration, motility, morphology, vitality	v	v	v	~	v	P	A	P	G	A	L
Lassen et al. 2014	General population, "young men" undergoing physical examination for military service	single spot urine, osmolality adjusted	Number, concentration, motility, morphology	~	~	v	~	nd	P	G	G	G	G	M/H
Li et al. 2011	Cohort, occupationally exposed	two spot urine samples, pre- and post-shift; for controls one spot sample; volume-based- and creatinine-corrected	Number, concentration, motility, morphology, vitality	v	v	v	~	v	A	G	G	G	G	M/H
Meeker et al. 2010	Cohort, male partners of subfertile couples 18 - 55 y	spot urine samples, various sampling schemes; some 2-3 samples; corrected for specific gravity	Number, concentration, motility, morphology	~	v	v	v	nd	A	G	A	G	G	M/H
Mendiola et al. 2010	Cohort, partners of pregnant women, general population	no details on urine samples, correction for creatinine	Number, concentration, motility, morphology	~	~	~	~	nd	P	G	A	G	G	M
Omran et al. 2018	Case-control, infertile men	single spot urine sample, creatinie adjusted	Concentration, motility, morphology	nd	~	~	~	nd	P	A	P	P/CD	A	U
Radwan et al. 2018	Cross-sectional; men attending fertility clinic, but with normal sperm parameters	single spot urine sample, creatinine adjusted	Number, concentration, motility, morphology	~	~	v	~	nd	P	G	A	G	G	M

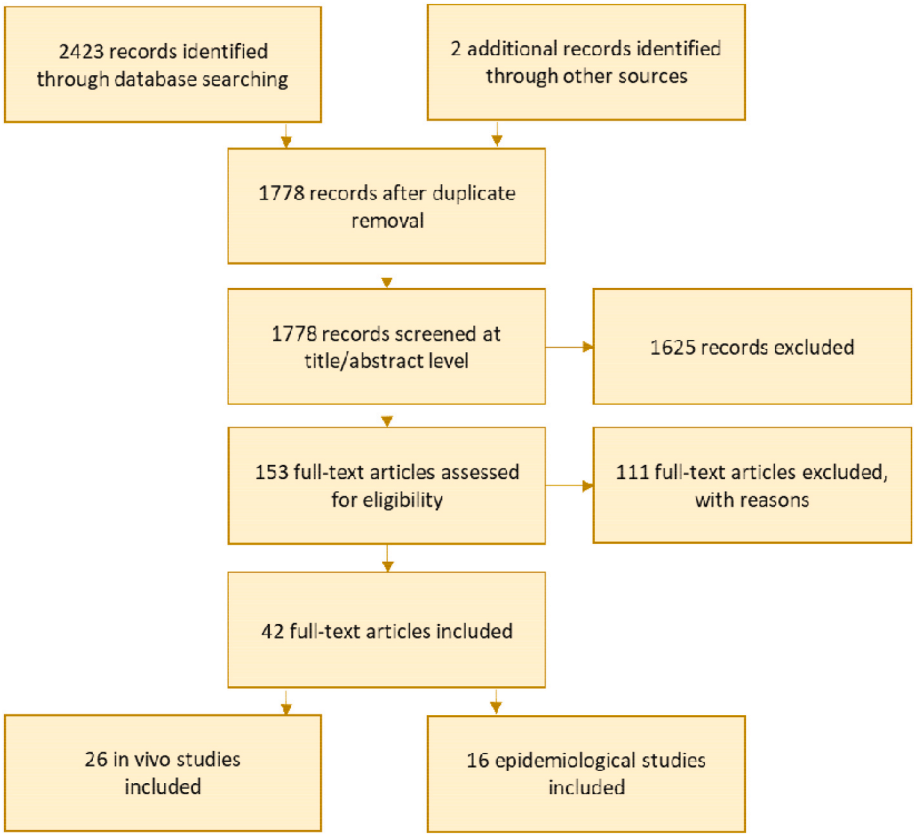


Fig. 1. Literature flow diagramme for animal studies and epidemiological studies of BPA exposures and semen quality.

framework developed by Radke et al. (2018) which assigns strength of evidence conclusions of *Robust*, *Moderate*, *Slight*, *Indeterminate*, and *Compelling evidence of no effect*. *Robust* describes evidence from *high* or *medium* confidence independent studies that report an association between BPA exposure and declines in semen quality, with reasonable confidence that alternative explanations, including chance, bias, and confounding, can be ruled out across studies. *Moderate* is used to describe a situation where there is a smaller number of studies (at least one *high* or *medium confidence* study with supporting evidence), with some heterogeneous results, that do not reach the degree of confidence required for *robust*. *Slight* is assigned when there are one or more studies reporting an association between bisphenol A and declining semen quality, but where considerable uncertainty exists. The evidence is limited to a set of consistent *low* confidence studies, or higher confidence studies with unexplained heterogeneity. *Indeterminate* is used when either there are no studies available in humans or when the evidence is highly inconsistent and primarily of *low* confidence. *Compelling evidence of no effect* requires several *high* confidence epidemiological studies returning null results.

2.5. Derivation of a BPA reference dose for declines in semen quality

To derive a bisphenol A reference dose with respect to declines in semen quality, we followed the procedure sketched out in EFSA (2017). Briefly, we made quantitative comparisons for each line of evidence (per animal species, and human) where it was possible to derive a point of departure (NOAEL or benchmark dose).

Where necessary, NOAELs were extrapolated from LOAELs by using a standard assessment factor (AF = 3). We based these comparisons on high quality studies (*high* or *medium* confidence human studies, *Tier 1* animal studies).

In humans, smaller doses than in animals are required to achieve the same effective tissue concentrations. To address species-specific differences in the toxicokinetics of BPA, we applied the points of departure identified (or extrapolated) from animal data to derive a human equivalent dose (HED) by application of human equivalent dose factors (HEDF), as detailed in EFSA (2021). The HEDs were then compared with BPA exposure ranges from epidemiological studies.

For comparisons of dosages used in animal studies with exposures experienced by humans, we converted urinary BPA levels to estimated daily intakes using the model developed by Koch et al. (2012).

3. Results

The literature selection process for human epidemiological studies and animal studies is shown in Fig. 1.

We first assessed the strength of evidence for an association between BPA exposure and declines in semen quality and then attempted to estimate a reference dose for this health endpoint for use in mixture risk assessments.

3.1. Strength of evidence: experimental studies in animals

Study selection and evaluation: Twenty-six experimental studies of BPA exposure and deteriorations of semen quality in rats and mice met our eligibility criteria (Table 1). Except for the study by Vom Saal et al. (1998), there were no concerns regarding the purity of the test compound, its consistent administration or possible contaminations, as BPA of a purity >99% was used in all studies.

Some studies, however, raised concerns about hormonal disturbances introduced through soy-containing diets. This is of relevance, as dosing of rats with genistein from GD 7 to the end of pregnancy led to declines in semen quality (Delclos et al., 2001) and feeding soy-containing diets obscured the effects of diethylstilboestrol on semen quality (Ruhlen et al., 2011). In several studies there was direct evidence that the diets used contained phytoestrogens. This was the case with

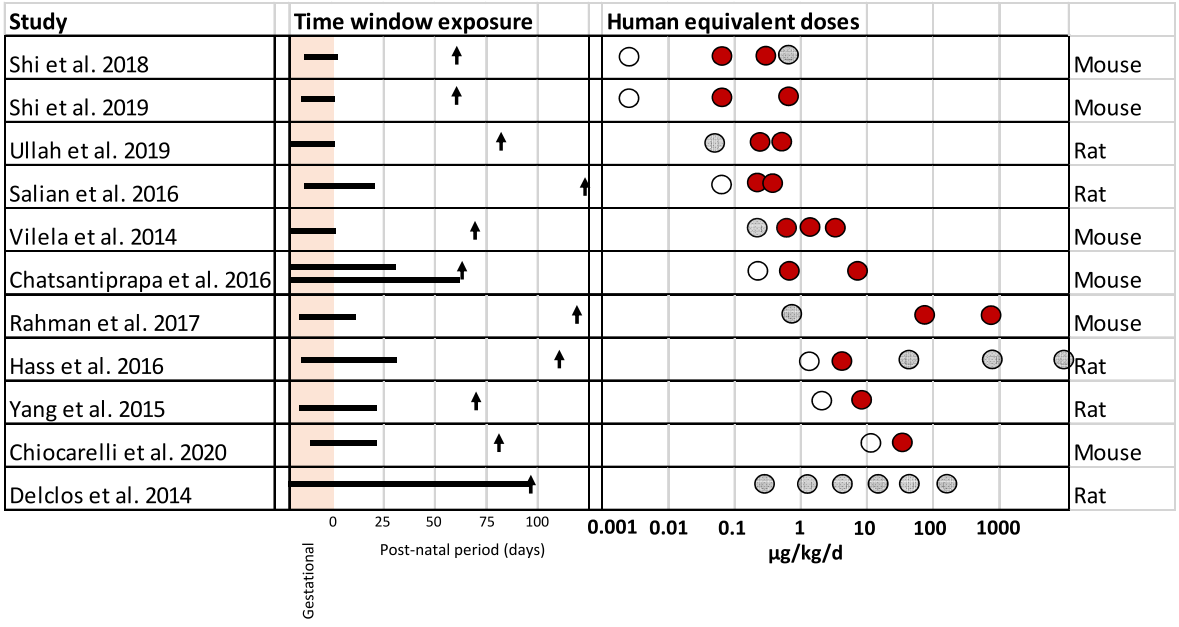


Fig. 2. Summary of high confidence (*Tier 1*) animal studies of BPA and semen quality. Black horizontal bars in "Time window exposure" show the periods of BPA administration, arrows depict time points when semen was sampled. The gestational period is shaded pink. Open circles in "Human equivalent doses" are doses equivalent to no-observed adverse effect levels (NOAELs) which were extrapolated from lowest-observed adverse effect levels (LOAELs) by application of an assessment factor of 3. Red circles are doses associated with declines in sperm numbers, grey circles show doses without effects on sperm numbers (if these are also the lowest tested doses, they are NOAELs). (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

Table 3
Estimation of reference doses for semen quality declines from animal studies.

	LOAEL (µg/kg d)	NOAEL (µg/kg d)	Species	HEDF	HED (µg/kg d)	RfD (µg/kg d)
Shi et al., 2018	0.5	**0.17**	Mouse	0.0155	0.0026	0.0001
Shi et al., 2019	0.5	**0.17**	Mouse	0.0155	0.0026	0.0001
Salian et al., 2009	1.2	**0.4**	Rat	0.165	0.0660	0.0026
Ullah et al., 2019	1.5	0.5	Rat	0.165	0.0825	0.0033
Vilela et al., 2014	40	**13.3**	Mouse	0.0155	0.2062	0.0082
Chatsantiprapa et al., 2016	50	**16**	Mouse	0.0155	0.2480	0.0099

LOAEL: Lowest observed adverse effect level; NOAEL: No-observed adverse effect level; HEDF: Human equivalent dose factor; HED: Human equivalent dose; RfD: Reference dose derived by dividing HED values by 25.
NOAEL values shown in bold are extrapolations from studies where only LOAELs, but not NOAELs were observed. In these cases, LOAELs were divided by 3 and the resulting values taken as "extrapolated" NOAELs.

Cagen et al. (1999) who used a rodent diet containing dehulled soybean meal. Rat chow 5001, 5002 and 5008 used by Howdeshell et al. (2008), Tyl et al. (2002, 2008) and Vom Saal et al. (1998) also contains phytoestrogens. The same applies to standard CE2 diet fed by Kobayashi et al. (2010, 2012) or standard chow by Meng et al. (2018). Accordingly, we rated these studies as low confidence (*Tier 3*) and did not conduct further detailed evaluations. Ema et al. (2001) did not report on the phytoestrogen content of the diet used, but we considered this as probably low risk because they observed endocrine-related BPA effects (changes in anogenital distance, but not semen parameters). The same applies to Chatsantiprapa et al. (2016), Shi et al. (2018, 2019) and Yang et al. (2015). Yoshino et al. (2002) also provided insufficient information on the phytoestrogen content and did not observe endocrine BPA effects. There were doubts about the proficiency of this study to demonstrate BPA effects on semen quality, as positive controls were not included. Accordingly, we evaluated Yoshino et al. (2002) as low confidence and assigned this study to *Tier 3*.

All remaining studies were evaluated as "definitely" or "probably low risk" in terms of randomisation, concealment and blinding and in relation to statistical power of detecting an effect (sufficient number of animals). Only Rahman et al. (2017) and Yang et al. (2015) employed fewer than 5 animals per dose group but possible concerns about insufficient power were made baseless by their observations of BPA-related effects on parameters of semen quality.

In addition to Yoshino et al. (2002), several other studies gave reason to doubt the sensitivity of the model used and its proficiency in detecting effects on semen quality. There was direct evidence for a lack of sensitivity in Kendig et al. (2017), Cagen et al. (1999) and Tyl et al. (2008) who employed ethinylestradiol, DES or estradiol, respectively, as a positive control but were unsuccessful in observing effects ("definitely high risk"). In all the other remaining cases (Dere et al., 2018; Ema et al., 2001; Nagao et al., 2002; Spörndly-Nees et al., 2018), there was indirect evidence for lack of sensitivity as positive controls were not used and BPA effects on semen parameters were not observed. This resulted in a rating of probably high risk and assignment to *Tier 3*.

With the remaining studies there were no concerns regarding BPA background contamination or any of the other evaluation elements. Chioccarelli et al. (2020) and Yang et al. (2015) did not provide adequate information about the statistical methods they used for estimating BPA doses associated with small effects, which we deemed to have an impact on study validity ("probably high risk"). However, this did not influence the overall confidence rating of "high" (*Tier 1*) for these studies.

Overall study confidence ratings: In summary, 11 of the 26 eligible

Author	Study design	Number of subjects	BPA urinary levels without effect (µg/g creatinine)	Estimated daily intakes (µg/kg d)
Adoamnei 2018	cross-sectional	215	1.3 - 2.8	**0.04 - 0.08**
Pollard 2019	cohort	161	2.39	**0.07**
Ji 2018	cross-sectional	500	0.32 - 1.12	**0.01 - 0.03**
Lassen 2014	cohort	297	3.71 - 6.44	**0.10 - 0.18**

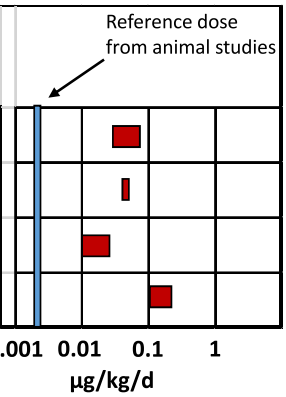


Fig. 3. Comparison of BPA reference dose with BPA exposure ranges not associated with semen quality declines in selected epidemiological studies. The red horizontal bars represent estimated daily BPA intakes for which effects on semen quality were not noted. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

studies obtained a confidence rating of "high" (*Tier 1*). In these studies, all six key evaluation elements were evaluated as "definitely" or "probably low risk", with no more than one other element rated as "definitely" or "probably high risk". The other 15 studies only achieved a confidence rating of "low" and accordingly had to be placed in *Tier 3*. Several of the *Tier 3* studies failed multiple key elements of the evaluation (Table 1). There were no *Tier 2* studies.

Evidence synthesis: As shown in Table 1, 10 of the 11 studies with a high confidence rating reported effects of BPA on semen quality parameters. The only high confidence study that did not observe effects was by Delclos et al. (2014).

Of the 15 studies we evaluated as being of low confidence, 5 reported BPA effects, while 10 did not observe declining semen quality after BPA exposure.

In summary, there are 10 independent studies with a "high" (*Tier 1*) confidence rating which reported declining semen quality after BPA exposure in multiple strains of two species, rat and mouse. Accordingly, the overall strength of evidence that associates BPA with poor semen quality in experimental studies can be evaluated as "robust".

3.2. Strength of evidence: human epidemiological studies

Study selection and evaluation: We identified 16 studies that matched our eligibility criteria (Table 2). These studies are case-control, cohort or cross-sectional with participants drawn from the general population, occupational cohorts, or couples from infertility clinics. The studies varied in size from 105 (Caporossi et al., 2020) to 1590 participants (Chen et al., 2013).

BPA measurements corresponding to prenatal exposures would be ideal for investigating associations with semen quality, as the *in utero* environment is critical for semen quality in adulthood (Skakkebaek et al., 2015). However, none of the eligible studies related semen quality to maternal exposures. The few studies that investigated exposures during development had to be excluded due to measurement of bisphenol A in serum, cord blood or seminal fluids for which toxicokinetic models for conversion to daily intakes are missing (Hart et al., 2018; Vitku et al., 2016).

In adult men, the best timing of exposure measurements would be around 90 days before taking a semen sample, because spermatogenesis takes approximately 75 days, with an additional 12 days of maturation as the sperm travels through the epididymis. However, none of the eligible studies adopted such a timing. Instead, most studies collected urine samples for BPA measurements at the same time, or near the time of semen analysis (Caporossi et al., 2020; Chen et al., 2013; Ji et al., 2018; Knez et al., 2014; Lassen et al., 2014). The exception is Pollard et al. (2019) who measured BPA exposures 3 days before sampling semen. Many studies did not give explicit details as to the timing of exposure measurements (Adoamnei et al., 2018; Benson et al., 2021; Den Hond et al., 2015; Goldstone et al., 2015; Kim et al., 2019; Li et al., 2011; Meeker et al., 2010; Mendiola et al., 2010; Omran et al., 2018; Radwan et al., 2018).

BPA has a relatively short half-life of excretion of 4–5 h. This may result in considerable variations of urinary BPA concentrations. To take account of these variations, sampling at multiple time points is recommended (Agier et al., 2020). However, most of the eligible studies based their BPA exposure estimates on single spot urine samples.

Due to the shortcomings regarding timing and frequency of urine sampling, we rated the exposure assessments in most of the studies as "poor". The only exceptions are Pollard et al. (2019), Li et al. (2011) and Meeker et al. (2010) who employed multiple BPA measurements which we regarded as somewhat mitigating the shortcomings regarding the timing of exposure measurements. Accordingly, the exposure assessments in these three studies were rated as "adequate".

We evaluated the outcome measurements as "good" when semen analyses were conducted according to the WHO (2010) guidelines. This applied to almost all studies, except when several of the core semen quality parameters (sample volume, sperm concentration, motility and morphology) were not examined, as in Chen et al. (2013) and Kim et al. (2019). We rated the outcome measurements in these two studies as "poor". Studies with missing descriptions of semen quality measurements were classed as critically deficient (Kim et al., 2019). In some cases, motility assessments could not be performed, as semen samples were collected at home and then shipped for analysis. We rated these studies as "adequate" (Pollard et al., 2019; Goldstone et al., 2015). Knez et al. (2014) and Li et al. (2011) also examined sperm vitality.

Studies that chose subjects from the general population, with no apparent selection effects and high participation rates were evaluated as "good" in relation to participant selection (Adoamnei et al., 2018; Benson et al., 2021; Goldstone et al., 2015; Ji et al., 2018; Lassen et al., 2014; Pollard et al., 2019). We classed occupational studies and those in infertility clinic settings as "adequate" (Caporossi et al., 2020; Chen et al., 2013; Den Hond et al., 2015; Meeker et al., 2010; Mendiola et al., 2010; Radwan et al., 2018). Where details on participant selection were missing, we applied the rating of "poor" (Kim et al., 2019; Omran et al., 2018).

Key confounders that must be considered in semen quality studies include age, abstinence time, smoking, body mass index, and chronic diseases (Sánchez-Pozo et al., 2013). Although not as well established as risk factors, alcohol use and stress may also warrant consideration. Most eligible studies adjusted for the key confounders, and accordingly we evaluated them as "good" in terms of confounder analysis. Where abstinence time was not included as a confounder or where information about abstinence time was missing or where subjects with abstinence times of fewer than 2 days were included in the analysis, we applied a rating of "poor" (Benson et al., 2021; Den Hond et al., 2015; Pollard

et al., 2019; Kim et al., 2019; Omran et al., 2018).

Ideally, evaluations of associations of chemical exposures with semen quality should analyse semen parameters as continuous variables, to minimise misclassification and to obtain sufficient statistical power. Furthermore, results should be presented with standard errors and confidence intervals and not just shown as "significant". Most of the studies met these requirements and were rated as "good" in terms of data analysis. Chen et al. (2013), Pollard et al. (2019) and Meeker et al. (2010) dichotomised semen quality parameters, and accordingly, we downgraded this evaluation aspect in these studies to "adequate". Kim et al. (2019) provided insufficient detail of their statistical analysis and had to be rated as "poor".

Overall study confidence ratings: Due to the importance of the exposure assessment component, we judged that no study with an exposure assessment rating of "poor" should obtain an overall confidence rating of "high". Studies where all other aspects were evaluated as "good" could achieve a maximum overall confidence rating of "medium to high" (M/H). If three or more components were evaluated as "poor", we applied an overall rating of "uninformative" (U). With two aspects classed as "poor", the overall rating was pegged at "low" (L). Table 2 shows the study confidence ratings we established according to these decision rules.

Evidence synthesis: As shown in Table 2, 8 studies returned null findings and 8 reported associations of declining semen parameters with BPA exposures. Of the 8 null studies, 3 achieved an overall confidence rating of "medium" (Caporossi et al., 2020; Goldstone et al., 2015; Mendiola et al., 2010) while the others were evaluated either as "low" or "uninformative". Among the 8 studies that found associations with BPA, 5 were "medium to high", and 2 "medium" and one study was rated "low" (Knez et al., 2014).

The disparity between the 3 "medium" confidence null studies and those that reported associations can be attributed to differences in exposure conditions: Caporossi et al. (2020), Goldstone et al. (2015) and Mendiola et al. (2010) all examined populations with rather low BPA urinary levels. This may well have precluded the detection of associations with BPA. Thus, rather than yielding conflicting evidence (unexplained positive and negative results in similarly exposed human populations) the eligible studies produced mixed results explained by differing exposure levels.

In summary, there are 7 independent studies of "medium" or "medium to high" confidence with positive findings. Accordingly, the overall strength of evidence of associations between BPA exposures and declines in semen quality can be evaluated as "robust".

3.3. Weight of evidence

There is robust evidence from animal studies that BPA exposures during gestation lead to declines in semen quality. In humans, evidence of the consequences of BPA exposures in fetal life is currently not available. However, the associations of BPA exposure in adult life with declines in semen quality are robust and support the conclusion that the patterns seen in animal experiments are relevant to humans. They are sufficiently robust to support hazard identification and characterisation. Accordingly, we proceeded to attempt a derivation of a BPA reference dose for declines in semen quality (hazard characterisation).

3.4. Derivation of a reference dose for declines in semen quality

Experimental studies in animals: Fig. 2 summarises all *Tier 1* studies with respect to the time windows of exposures used and the BPA doses associated with statistically significant declines in semen quality (sperm numbers and motility). All studies covered the critical period when germ cell stem populations are established (mouse: GD 7 to PND 8, rat: GD 9 to PND 10).

Due to higher rates of metabolism and excretion in rodents, the doses required to attain comparable tissue levels in mice, rats or humans differ. Normally, higher doses than in humans are required to achieve similar tissue levels in rodents. The availability of serum-concentration time course data allows making such comparisons on a quantitative basis, in terms of Areas under the Curve (AUC). To adjust for kinetic differences, and to make the exposures comparable, AUCs resulting from comparable doses in animal species are divided by AUCs in humans to obtain Human Equivalent Dose Factors (HEDF). Human Equivalent Doses (HED) are then obtained by multiplying the doses used in rodent studies with the appropriate HEDF (0.0155 for mouse and 0.165 for rat studies, respectively) (EFSA, 2021). We focused on studies with at least two different dose groups, in addition to untreated controls. As shown in Fig. 2, the lowest doses associated with declines in sperm numbers varied from a HED of 0.0077 μg/kg/d (Shi et al., 2018, 2019) to 77.5 μg/kg/d (Rahman et al., 2017), with most studies reporting activity at HEDs between 0.24 and 8.25 μg/kg/d.

In estimating a BPA reference dose, we first calculated HEDs based on no-observed adverse effect levels (NOAELs). In almost all studies, the lowest used treatment doses produced effects which precluded the determination of a NOAEL. In these cases, we extrapolated NOAELs from the reported lowest treatment doses (lowest observed adverse effect levels, LOAELs) by application of an AF of 3. HEDs from studies with observed or extrapolated NOAELs above 1 μg/kg/d were not considered further. To account for species differences and vulnerable individuals, we adopted the procedure described in EFSA (2015, 2021) and divided the HEDs by an assessment factor of 25, widely used by EFSA for chemical risk assessment. This produced the reference doses listed in Table 3. The only study that reported a NOAEL was by Ullah et al. (2019). Their NOAEL was 0.5 μg/kg/d in the rat, which by combination with an HEDF of 0.165 and an AF of 25 produces a reference dose of 0.0033 μg/kg/d. A very similar reference dose of 0.0026 μg/kg/d was estimated based on the data in Salian et al. (2009). Based on the findings by Shi et al. (2018, 2019) in mice, we estimated 0.0001 μg/kg/d, and the mouse studies by Vilela et al. (2014) and Chatsantiprapa et al. (2016) produced 0.0082 and 0.0099 μg/kg/d, respectively. It appears that the estimates derived from Ullah et al. (2019) and Salian et al. (2009) occupy a mid-point, and accordingly, we adopted 0.003 μg/kg/d as a BPA reference dose in mixture risk assessments for declines in semen quality.

Comparison of BPA reference dose with data from human epidemiological studies: We compared the reference dose estimate derived from animal studies with BPA exposures in epidemiological studies below the ranges associated with declines in semen quality ("no-observed effect ranges"). We based our comparison on studies among the general population and excluded occupationally exposed cohorts and populations from fertility clinics. This left four studies eligible for this comparison: Adoamnei et al. (2018), Pollard et al. (2019), Ji et al. (2018) and Lassen et al. (2014) (Fig. 3).

The authors of these studies categorised BPA exposures into ranges of urinary concentrations which they analysed in terms of statistically significant associations with declines in semen quality. We converted the urinary BPA concentrations reported in these studies into estimated daily intakes, by using the model developed by Koch et al. (2012). This allowed us to identify exposure ranges apparently no longer associated with semen quality. As shown in Fig. 3, the reference dose of 0.003 μg/kg/d estimated from animal studies is below the "no-observed effect range" of between 0.01 and 0.18 μg/kg/d reported in the four epidemiological studies (Fig. 3).

3.5. Comparison with BPA exposure estimates

Using the model by Koch et al. (2012), the urinary BPA concentrations from samples collected in 2009 in human biomonitoring exercises by Koch et al. (2012) and Frederiksen et al. (2020) were converted into estimated daily intakes of 0.14 and 0.16 μg/kg/d, respectively (95th percentiles). The median daily intakes were reported as 0.035 and 0.048 μg/kg/d, respectively. These estimates from German and Danish

subjects agree well with those for Norwegian populations published by Karrer et al. (2020) (median: 0.035 μg/kg/d, most probable range: 0.02–0.1 μg/kg/d).

Thus, European populations experience BPA exposures 12 to 16-fold (median) and up to 48-fold (95th percentile) above the reference dose we estimated from animal studies. These exposures fall in the ranges where declines in semen quality were observed in epidemiological studies (Adoamnei et al., 2018: 0.045–0.08 μg/kg/d; Pollard et al., 2019: 0.07 μg/kg/d; Ji et al., 2018: 0.01–0.03 μg/kg/d; Lassen et al., 2014: 0.1–0.18 μg/kg/d).

4. Discussion

The evidence linking BPA exposures to declines in semen quality is often characterised as "conflicting" or "varied". In contrast with such views, the application of the systematic review method, together with a rigorous confidence rating approach, reveals that there is convincing evidence of declines of semen quality after gestational BPA exposures in animal studies. In addition, human epidemiological studies provide supporting evidence for poor semen quality after BPA exposure in adult life. This is despite the noted weaknesses in exposure assessments which will have increased the likelihood of null findings through exposure misclassification (Agier et al., 2020).

We propose that a great deal of the negative findings in animal studies can be attributed to deficiencies in study sensitivity and to insufficient control of background contamination with BPA due to the use of polycarbonate caging. Howdeshell et al. (2003) demonstrated that BPA leaches from polycarbonate cages. In prepubertal mice housed in such cages, Howdeshell et al. saw increases in uterine weights (albeit not statistically significant). Some studies also raised concerns about confounding through hormonal interference by phytoestrogens from soy-containing diets. Ruhlen et al. (2011) observed that soy-containing diets obscure the effects of diethylstilboestrol on semen quality. Furthermore, gestational exposures of rats to phytoestrogens such as genistein led to declines in semen quality (Delclos et al., 2001).

Apart from general concerns about the quality of exposure assessments, the confidence in some epidemiological studies was compromised by deficiencies in outcome measurements and adjustments for confounding. Nevertheless, the inconsistent findings from "medium to high" and "medium" confidence studies can be explained in terms of differences in exposure conditions. This is to be distinguished from "conflicting evidence" in the sense of conflicting findings due to unexplained factors.

On this robust basis, we attempted the estimation of BPA exposures very likely not associated with declines in semen quality. This estimate is intended for use in future mixture risk assessments of male reproductive health.

Our value is derived from the data in animal studies. For most "high" confidence studies (*Tier 1*), the HEDs calculated from the corresponding LOAELs or NOAELs fall in the range between around 0.0026 and 0.25 μg/kg/d (Fig. 2, Table 3). By application of an AF of 25, these HEDs translate into 0.0001–0.01 μg/kg/d as possible BPA reference doses. Our choice of 0.003 μg/kg/d approximates the data from Salian et al. (2009) and Ullah et al. (2019). We judged that the higher estimates of 0.008 and 0.01 μg/kg/d which could have been chosen based on Vilela et al. (2014) and Chatsantiprapa et al. (2016), respectively, would have been insufficiently conservative, considering that Shi et al. (2018, 2019) reported effects at approximately 100-fold lower doses.

The reason why we did not opt for 0.0001 μg/kg/d, as supported by Shi et al. (2018, 2019), lies in the purpose of this exercise. Rather than providing a high degree of protection, as is essential when deriving health-based guidance values or TDIs, the intended use of our value in a mixture risk assessment dictated our interest in a *reasonable estimate* of BPA exposures likely without effects on semen quality. This led us to weigh the low doses in Shi et al. (2018, 2019) against the higher levels observed in the other animal studies. We realise that this procedure does not conform with the standards necessary for deriving tolerable daily intakes for single chemical exposures. We would like to emphasise that our business here is *not* in deriving a health-based guidance value or TDI for BPA which indeed would require a higher degree of conservatism and perhaps a correspondingly lower reference dose.

Our quantitative comparison of dose ranges in animal studies with exposure levels in epidemiological studies was for orientation only and did not influence our choice of a BPA reference dose. We recognise that such comparisons are problematic as the human studies related semen quality to contemporaneous BPA exposures, not gestational exposures. To our knowledge, epidemiological studies of gestational BPA exposures are not available. However, there is evidence from animal studies that BPA exposure in adulthood also leads to poor semen quality (not reviewed here, but for examples see Wang et al., 2016; Ullah et al., 2018).

While our evaluations were in progress, the EFSA Panel on Food Contact Materials, Enzymes, Flavourings and Processing Aids (CEF) completed their re-evaluation of the 2015 temporary TDI for BPA. They proposed a new TDI of 0.04 ng/kg body weight/day (EFSA 2021), 100,000-times lower than the previous value of 4 μg/kg body weight/day (EFSA 2015). This new estimate considers immunotoxic effects as critical. Of relevance to our assessment, the EFSA Panel considered BPA effects on semen parameters that result from gestational or post-natal exposures until weaning as unlikely, based on animal studies that appeared up to 2018. However, in their assessment, EFSA did not consider several studies we rated as high confidence (Tier 1), such as Chatsantiprapa et al. (2016), Vom Saal et al. (1998), Vilela et al. (2014), and Yang et al. (2015). Shi et al. (2019) and Ullah et al. (2019) were published outside EFSA's evaluation period. Thus, the two studies that most heavily influenced our estimate, Salian et al. (2009) and Ullah et al. (2019), did not find entry into EFSA's evaluation. Our assessment agrees with the appraisal of studies in a bisphenol S evaluation by Beausoleil et al. (2022) who regarded Shi et al. (2019) and Ullah et al. (2019) as key studies.

In contrast to their view of the strength of evidence from gestational and post-natal animal studies, EFSA judged effects of BPA exposures on semen quality (motility and viability) in adulthood as likely. This appraisal is based on the study by Wang et al. (2016) in adult mice who reported a LOAEL of 10 μg/kg/d for BPA effects on sperm motility. With an AF = 3, this gives an extrapolated NOAEL of 3.3 μg/kg/d, in good agreement with the lower limit benchmark dose of 3.41 μg/kg/d calculated by the EFSA Panel. Combined with a mouse HEDF of 0.0155 and a further AF of 25, this produces a reference dose of 0.0015 μg/kg/d, two-fold lower than our estimate of 0.003 μg/kg/d.

We intend to utilise the reference dose derived here in a mixture risk assessment for male reproductive health, with a focus on declines in semen quality. This will be of importance in view of the reported declines in semen quality, mainly in Western countries (Levine et al., 2017). The assessment will include multiple chemicals, such as polychlorinated dibenzo-dioxins, polychlorinated biphenyls, phthalates, bisphenol F and S, parabens and many more. It is hoped that this will bring the contours of chemical exposures that impact on fertility into view.

Declaration of competing interest

The authors declare there are no conflicts of interest.

Acknowledgements

This work was conducted with funding from the European HBM4EU project (www.hbm4eu.eu), contract number 733032, Horizon 2020 programme, which is gratefully acknowledged.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.ijheh.2022.113942.

References

Adoamnei, E., Mendiola, J., Vela-Soria, F., Fernández, M.F., Olea, N., Jørgensen, N., Swan, S.H., Torres-Cantero, A.M., 2018. Urinary bisphenol A concentrations are associated with reproductive parameters in young men. Environ. Res. 161, 122–128. https://doi.org/10.1016/j.envres.2017.11.002.

Agier, L., Slama, R., Basagaña, X., 2020. Relying on repeated biospecimens to reduce the effects of classical-type exposure measurement error in studies linking the exposome to health. Environ. Res. 186, 109492. https://doi.org/10.1016/j.envres.2020.109492.

Apel, P., Kortenkamp, A., Koch, H.M., Vogel, N., Rüther, M., Kasper-sonnenberg, M., Conrad, A., Brüning, T., Kolossa-gehring, M., 2020. Time course of phthalate cumulative risks to male developmental health over a 27-year period : biomonitoring samples of the German Environmental Specimen Bank. Environ. Int. 137, 105467. https://doi.org/10.1016/j.envint.2020.105467.

Axelstad, M., Christiansen, S., Boberg, J., Scholze, M., Jacobsen, P.R., Isling, L.K., Kortenkamp, A., Hass, U., 2014. Mixtures of endocrine-disrupting contaminants induce adverse developmental effects in preweaning rats. Reproduction 147, 489–501. https://doi.org/10.1530/REP-13-0447.

Axelstad, M., Hass, U., Scholze, M., Christiansen, S., Kortenkamp, A., Boberg, J., 2018. EDC impact: reduced sperm counts in rats exposed to human relevant mixtures of endocrine disrupters. Endocr. Connect. 7, 139–148. https://doi.org/10.1530/EC-17-0307.

Bauer, A.Z., Swan, S.H., Kriebel, D., Liew, Z., Taylor, H.S., Bornehag, C.G., Andrade, A. M., Olsen, J., Jensen, R.H., Mitchell, R.T., Skakkebaek, N.E., Jégou, B., Kristensen, D. M., 2021. Paracetamol use during pregnancy — a call for precautionary action. Nat. Rev. Endocrinol., 0123456789 https://doi.org/10.1038/s41574-021-00553-7.

Beausoleil, C., Le Magueresse-Battistoni, B., Viguié, C., Babajko, S., Canivenc-Lavier, M. C., Chevalier, N., Emond, C., Habert, R., Picard-Hagen, N., Mhaouty-Kodja, S., 2022. Regulatory and academic studies to derive reference values for human health: the case of bisphenol S. Environ. Res. 204 https://doi.org/10.1016/j.envres.2021.112233.

Benson, T.E., Gaml-Sørensen, A., Ernst, A., Brix, N., Hougaard, K.S., Hærvig, K.K., Bonde, J.P.E., Tøttenborg, S.S., Lindh, C.H., Ramlau-Hansen, C.H., Toft, G., 2021. Urinary bisphenol a, f and s levels and semen quality in young adult Danish men. Int. J. Environ. Res. Publ. Health 18, 1–12. https://doi.org/10.3390/ijerph18041742.

Cagen, S.Z., Waechter, J.M., Dimond, S.S., Breslin, W.J., Butala, J.H., Jekat, F.W., Joiner, R.L., Shiotsuka, R.N., Veenstra, G.E., Harris, L.R., 1999. Normal reproductive organ development in wistar rats exposed to bisphenol A in the drinking water. Regul. Toxicol. Pharmacol. 30, 130–139. https://doi.org/10.1006/rtph.1999.1340.

Caporossi, L., Alteri, A., Campo, G., Paci, E., Tranfo, G., Capanna, S., Papaleo, E., Pigini, D., Viganò, P., Papaleo, B., 2020. Cross sectional study on exposure to BPA and phthalates and semen parameters in men attending a fertility center. Int. J. Environ. Res. Publ. Health 17. https://doi.org/10.3390/ijerph17020489.

Chatsantiprapa, K., Sophon, T., Sattayasai, J., 2016. Effects of continuous exposure to bisphenol A on male and female mice from prenatally to adulthood. Thai Journal of Pharmceutical Sciences 40 (2), 61–69.

Chen, M., Tang, R., Fu, G., Xu, Bin, Zhu, P., Qiao, S., Chen, X., Xu, Bo, Qin, Y., Lu, C., Hang, B., Xia, Y., Wang, X., 2013. Association of exposure to phenols and idiopathic male infertility. J. Hazard Mater. 250–251, 115–121. https://doi.org/10.1016/j.jhazmat.2013.01.061.

Chioccarelli, T., Manfrevola, F., Migliaccio, M., Altucci, L., Porreca, V., Fasano, S., Cobellis, G., Cescon, M., 2020. Fetal-perinatal exposure to bisphenol-A affects quality of spermatozoa in adulthood mouse. Internet J. Endocrinol. https://doi.org/10.1155/2020/2750501, 2020.

Christiansen, S., Axelstad, M., Boberg, J., Vinggaard, A.M., Pedersen, G.A., Hass, U., 2014. Low-dose effects of Bisphenol a on early sexual development in male and female rats. Reproduction 147, 477–487. https://doi.org/10.1530/REP-13-0377.

De Gendt, K., Swinnen, J.V., Saunders, P.T.K., Schoonjans, L., Dewerchin, M., Devos, A., Tan, K., Atanassova, N., Claessens, F., Lécureuil, C., Heyns, W., Carmeliet, P., Guillou, F., Sharpe, R.M., Verhoeven, G., 2004. A Sertoli cell-selective knockout of the androgen receptor causes spermatogenic arrest in meiosis. Proc. Natl. Acad. Sci. U. S. A. 101, 1327–1332. https://doi.org/10.1073/pnas.0308114100.

de Rooij, D.G., Vergouwen, R.P., 1991. The estimation of damage to testicular cell lineages. Prog. Clin. Biol. Res. 372, 467–480.

Delclos, K.B., Bucci, T.J., Lomax, L.G., Latendresse, J.R., Warbritton, A., Weis, C.C., Newbold, R.R., 2001. Effects of dietary genistein exposure during development on male and female CD (Sprague-Dawley) rats. Reprod. Toxicol. 15, 647–663. https://doi.org/10.1016/S0890-6238(01)00177-0.

Delclos, K.B., Camacho, L., Lewis, S.M., Vanlandingham, M.M., Latendresse, J.R., Olson, G.R., Davis, K.J., Patton, R.E., Da costa, G.G., Woodling, K.A., Bryant, M.S., Chidambaram, M., Trbojevich, R., Juliar, B.E., Felton, R.P., Thorn, B.T., 2014. Toxicity evaluation of bisphenol a administered by gavage to sprague dawley rats from gestation day 6 through postnatal day 90. Toxicol. Sci. 139, 174–197. https://doi.org/10.1093/toxsci/kfu022.

Den Hond, E., Tournaye, H., De Sutter, P., Ombelet, W., Baeyens, W., Covaci, A., Cox, B., Nawrot, T.S., Van Larebeke, N., D'Hooghe, T., 2015. Human exposure to endocrine disrupting chemicals and fertility: a case-control study in male subfertility patients. Environ. Int. 84, 154–160. https://doi.org/10.1016/j.envint.2015.07.017.

Dere, E., Anderson, L.M., Huse, S.M., Spade, D.J., McDonnell-Clark, E., Madnick, S.J., Hall, S.J., Camacho, L., Lewis, S.M., Vanlandingham, M.M., Boekelheide, K., 2018. Effects of continuous bisphenol A exposure from early gestation on 90 day old rat testes function and sperm molecular profiles: a CLARITY-BPA consortium study. Toxicol. Appl. Pharmacol. 347, 1–9. https://doi.org/10.1016/j.taap.2018.03.021.

ECHA, (European Chemicals Agency), 2018. Seven new substances added to the Candidate List, entry for bisphenol-A updated [WWW Document], 11.11.21.

EFSA, 2015. Scientific Opinion on the risks to public health related to the presence of bisphenol A (BPA) in foodstuffs. PART II - Toxicological assessment and' 13 (1) https://doi.org/10.2903/j.efsa.2015.3978.

EFSA, Hardy, A., Benford, D., Halldorsson, T., Jeger, M.J., Knutsen, H.K., More, S., Naegeli, H., Noteborn, H., Ockleford, C., Ricci, A., Rychen, G., Schlatter, J.R., Silano, V., Solecki, R., Turck, D., Benfenati, E., Chaudhry, Q.M., Craig, P., Frampton, G., Greiner, M., Hart, A., Hogstrand, C., Lambre, C., Luttik, R., Makowski, D., Siani, A., Wahlstroem, H., Aguilera, J., Dorne, J., Fernandez Dumont, A., Hempen, M., Valtueña Martínez, S., Martino, L., Smeraldi, C., Terron, A., Georgiadis, N., Younes, M., 2017. Guidance on the use of the weight of evidence approach in scientific assessments. EFSA J. 15 (8) https://doi.org/10.2903/j.efsa.2017.4971.

EFSA, 2019. Testing the study appraisal methodology from the 2017 Bisphenol A (BPA) hazard assessment protocol. EFSA Support. Publ. 16. https://doi.org/10.2903/sp.efsa.2019.en-1732. (Accessed 8 February 2022).

EFSA, 2021. Re-evaluation of the risks to public health related to the presence of bisphenol A (BPA) in foodstuffs. Panel on Food Contact Material, Enzymes and Processing Aids (CEP). Scientific Opinion endorsed for public consultation. https://connect.efsa.europa.eu/RM/s/publicconsultation2/a0l1v00000E8BRD/pc0109.

Ema, M., Fujii, S., Furukawa, M., Kiguchi, M., Ikka, T., Harazono, A., 2001. Rat two-generation reproductive toxicity study of bisphenol A. Reprod. Toxicol. 15, 505–523. https://doi.org/10.1016/S0890-6238(01)00160-5.

Ermler, S., Scholze, M., Kortenkamp, A., 2011. The suitability of concentration addition for predicting the effects of multi-component mixtures of up to 17 anti-androgens with varied structural features in an in vitro AR antagonist assay. Toxicol. Appl. Pharmacol. 257, 189–197.

Frederiksen, H., Nielsen, O., Koch, H.M., Skakkebaek, N.E., Juul, A., Jørgensen, N., Andersson, A.M., 2020. Changes in urinary excretion of phthalates, phthalate substitutes, bisphenols and other polychlorinated and phenolic substances in young Danish men; 2009–2017. Int. J. Hyg Environ. Health 223, 93–105. https://doi.org/10.1016/j.ijheh.2019.10.002.

Goldstone, A.E., Chen, Z., Perry, M.J., Kannan, K., Louis, G.M.B., 2015. Urinary bisphenol A and semen quality, the LIFE study. Reprod. Toxicol. 51, 7–13. https://doi.org/10.1016/j.reprotox.2014.11.003.

Hart, R.J., Doherty, D.A., Keelan, J.A., Minaee, N.S., Thorstensen, E.B., Dickinson, J.E., Pennell, C.E., Newnham, J.P., McLachlan, R., Norman, R.J., Handelsman, D.J., 2018. The impact of antenatal Bisphenol A exposure on male reproductive function at 20–22 years of age. Reprod. Biomed. Online 36, 340–347. https://doi.org/10.1016/j.rbmo.2017.11.009.

Hass, U., Christiansen, S., Boberg, J., Rasmussen, M.G., Mandrup, K., Axelstad, M., 2016. Low-dose effect of developmental bisphenol A exposure on sperm count and behaviour in rats. Andrology 4, 594–607. https://doi.org/10.1111/andr.12176.

Howdeshell, K.L., Peterman, P.H., Judy, B.M., Taylor, J.A., Orazio, C.E., Ruhlen, R.L., vom Saal, F.S., Welshons, W.V., 2003. Bisphenol A is released from used polycarbonate animal cages into water at room temperature. Environ. Health Perspect. 111, 1180–1187. https://doi.org/10.1289/ehp.5993.

Howdeshell, K.L., Furr, J., Lambright, C.R., Wilson, V.S., Ryan, B.C., Gray, L.E., 2008. Gestational and lactational exposure to ethinyl estradiol, but not bisphenol a, decreases androgen-dependent reproductive organ weights and epididymal sperm abundance in the male long evans hooded rat. Toxicol. Sci. 102, 371–382. https://doi.org/10.1093/toxsci/kfm306.

Ji, H., Miao, M., Liang, H., Shi, H., Ruan, D., Li, Y., Wang, J., Yuan, W., 2018. Exposure of environmental Bisphenol A in relation to routine sperm parameters and sperm movement characteristics among fertile men. Sci. Rep. 8, 1–9. https://doi.org/10.1038/s41598-018-35787-5.

Karrer, C., Andreassen, M., von Goetz, N., Sonnet, F., Sakhi, A.K., Hungerbühler, K., Dirven, H., Husøy, T., 2020. The EuroMix human biomonitoring study: source-to-dose modeling of cumulative and aggregate exposure for the bisphenols BPA, BPS, and BPF and comparison with measured urinary levels. Environ. Int. 136, 105397. https://doi.org/10.1016/j.envint.2019.105397.

Kendig, E.L., Buesing, D.R., Christie, S.M., Cookman, C.J., Robin, B., Hugo, E.R., Kasper, S.N., Kendziorski, J.A., Ungi, K.R., 2017. Estrogen-like disruptive effects of dietary exposure to bisphenol A or 17α-ethinyl estradiol in CD1 mice. https://doi.org/10.1177/1091581812463254.Estrogen-Like, 31, 537-530.

Kim, H.K., Ko, D.H., Lee, W., Kim, K.R., Chun, S., Song, J., Min, W.K., 2019. Body fluid concentrations of bisphenol A and their association with in vitro fertilization outcomes. Hum. Fertil. 1–9. https://doi.org/10.1080/14647273.2019.1612104.

Knez, J., Kranvogl, R., Breznik, B.P., Vončina, E., Vlaisavljević, V., 2014. Are urinary bisphenol A levels in men related to semen quality and embryo development after medically assisted reproduction? Fertil. Steril. 101 https://doi.org/10.1016/j.fertnstert.2013.09.030.

Kobayashi, K., Kubota, H., Ohtani, K., Hojo, R., Miyagawa, M., 2012. Lack of effects for dietary exposure of bisphenol A during in utero and lactational periods on reproductive development in rat offspring. J. Toxicol. Sci. 37, 565–573. https://doi.org/10.2131/jts.37.565.

Kobayashi, K., Ohtani, K., Kubota, H., Miyagawa, M., 2010. Dietary exposure to low doses of bisphenol A: effects on reproduction and development in two generations of C57BL/6J mice. Congenital. Anom. 50, 159–170. https://doi.org/10.1111/j.1741-4520.2010.00279.x.

Koch, H.M., Kolossa-Gehring, M., Schröter-Kermani, C., Angerer, J., Brüning, T., 2012. Bisphenol A in 24 h urine and plasma samples of the German Environmental Specimen Bank from 1995 to 2009: a retrospective exposure evaluation. J. Expo. Sci. Environ. Epidemiol. 22, 610–616. https://doi.org/10.1038/jes.2012.39.

Kortenkamp, A., 2020. Molecular and Cellular Endocrinology Which chemicals should be grouped together for mixture risk assessments of male reproductive disorders. Mol. Cell. Endocrinol. 499, 110581. https://doi.org/10.1016/j.mce.2019.110581.

Lassen, T.H., Frederiksen, H., Jensen, T.K., Petersen, J.H., Joensen, U.N., Main, K.M., Skakkebaek, N.E., Juul, A., Jørgensen, N., Andersson, A.M., 2014. Urinary bisphenol a levels in young men: association with reproductive hormones and semen quality. Environ. Health Perspect. 122, 478–484. https://doi.org/10.1289/ehp.1307309.

Levine, H., Jørgensen, N., Martino-Andrade, A., Mendiola, J., Weksler-Derri, D., Mindlis, I., Pinotti, R., Swan, S.H., 2017. Temporal trends in sperm count: a systematic review and meta-regression analysis. Hum. Reprod. Update 23, 646–659. https://doi.org/10.1093/humupd/dmx022.

Li, D.K., Zhou, Z., Miao, M., He, Y., Wang, J., Ferber, J., Herrinton, L.J., Gao, E., Yuan, W., 2011. Urine bisphenol-A (BPA) level in relation to semen quality. Fertil. Steril. 95, 625–630. https://doi.org/10.1016/j.fertnstert.2010.09.026 e4.

Martin, O.V., Evans, R.M., Faust, M., Kortenkamp, A., 2017. A human mixture risk assessment for neurodevelopmental toxicity associated with polybrominated diphenyl ethers used as flame retardants. Environ. Health Perspect. 125 https://doi.org/10.1289/EHP826.

Martin, O.V., Baig, A., Ermler, S., McPhie, J., Scholze, M., Kortenkamp, A., 2021. Protocol for a systematic review of associations of bisphenol A exposure with declining semen quality in males to support derivation of a reference dose for mixture risk assessments for male reproductive health. Zenodo. https://doi.org/10.5281/zenodo.5083147.

Meeker, J.D., Ehrlich, S., Toth, T.L., Wright, D.L., Calafat, A.M., Trisini, A.T., Ye, X., Hauser, R., 2010. Semen quality and sperm DNA damage in relation to urinary bisphenol A among men from an infertility clinic. Reprod. Toxicol. 30, 532–539. https://doi.org/10.1016/j.reprotox.2010.07.005.

Mendiola, J., Jørgensen, N., Andersson, A.M., Calafat, A.M., Ye, X., Redmon, J.B., Drobnis, E.Z., Wang, C., Sparks, A., Thurston, S.W., Liu, F., Swan, S.H., 2010. Are environmental levels of bisphenol A associated with reproductive function in fertile men? Environ. Health Perspect. 118, 1286–1291. https://doi.org/10.1289/ehp.1002037.

Meng, Y., Lin, R., Wu, F., Sun, Q., Jia, L., 2018. Decreased capacity for sperm production induced by perinatal bisphenol A exposure is associated with an increased inflammatory response in the offspring of C57BL/6 male mice 1–11. https://doi.org/10.3390/ijerph15102158.

Moos, R.K., Apel, P., Schröter-Kermani, C., Kolossa-Gehring, M., Brüning, T., Koch, H.M., 2017. Daily intake and hazard index of parabens based upon 24h urine samples of the German environmental specimen bank from 1995 to 2012. J. Expo. Sci. Environ. Epidemiol. 27, 591–600. https://doi.org/10.1038/jes.2016.65.

Nagao, T., Saito, Y., Usumi, K., Yoshimura, S., Ono, H., 2002. Low-dose bisphenol A does not affect reproductive organs in estrogen-sensitive C57BL/6N mice exposed at the sexually mature, juvenile, or embryonic stage. Reprod. Toxicol. 16, 123–130. https://doi.org/10.1016/S0890-6238(02)00003-5.

NTP, 2015. Handbook for Conducting a Literature-Based Health Assessment Using OHAT Approach for Systemic Review and Evidence Integration 1–98.

NTP, 2019 Handbook for Conducting a Literature-Based Health Assessment Using OHAT Approach for Systematic Review and Evidence Integration, https://ntp.niehs.nih.gov/ntp/ohat/pubs/handbookmarch2019_508.pdf, accessed 8February 2022.

Olaso, R., Habert, R., 2000. Genetic and cellular analysis of male germ cell development. J. Androl. 21, 497–511.

Omran, G.A., Gaber, H.D., Mostafa, N.A.M., Abdel-Gaber, R.M., Salah, E.A., 2018. Potential hazards of bisphenol A exposure to semen quality and sperm DNA integrity among infertile men. Reprod. Toxicol. 81, 188–195. https://doi.org/10.1016/j.reprotox.2018.08.010.

Orton, F., Ermler, S., Kugathas, S., Rosivatz, E., Scholze, M., Kortenkamp, A., 2014. Mixture effects at very low doses with combinations of anti-androgenic pesticides, antioxidants, industrial pollutant and chemicals used in personal care products. Toxicol. Appl. Pharmacol. 278, 201–208. https://doi.org/10.1016/j.taap.2013.09.008.

Pollard, S.H., Cox, K.J., Blackburn, B.E., Wilkins, D.G., Carrell, D.T., Stanford, J.B., Porucznik, C.A., 2019. Male exposure to bisphenol A (BPA) and semen quality in the home observation of periconceptional exposures (HOPE) cohort. Reprod. Toxicol. 90, 82–87. https://doi.org/10.1016/j.reprotox.2019.08.014.

Radke, E.G., Braun, J.M., Meeker, J.D., Cooper, G.S., 2018. Phthalate exposure and male reproductive outcomes: a systematic review of the human epidemiological evidence. Environ. Int. 121, 764–793. https://doi.org/10.1016/j.envint.2018.07.029.

Radwan, M., Wielgomas, B., Dziewirska, E., Radwan, P., Kałużny, P., Klimowska, A., Hanke, W., Jurewicz, J., 2018. Urinary bisphenol A levels and male fertility. Am. J. Men's Health 12, 2144–2151. https://doi.org/10.1177/1557988318799163.

Rahman, M.S., Kwon, W.S., Karmakar, P.C., Yoon, S.J., Ryu, B.Y., Pang, M.G., 2017. Gestational exposure to bisphenol A affects the function and proteome profile of F1 spermatozoa in adult mice. Environ. Health Perspect. 125, 238–245. https://doi.org/10.1289/EHP378.

Ruhlen, R.L., Taylor, J.A., Mao, J., Kirkpatrick, J., Welshons, W.V., Vom Saal, F.S., 2011. Choice of animal feed can alter fetal steroid levels and mask developmental effects of endocrine disrupting chemicals. J. Dev. Orig. Health Dis. 2, 36–48. https://doi.org/10.1017/S2040174410000711.

Salian, Smita, Doshi, Tanvi, Vanage, Geeta, 2009. Perinatal exposure of rats to bisphenol A affects the fertility of male offspring. Life Sci. 85, 742–752.

Sánchez-Pozo, M.C., Mendiola, J., Serrano, M., Mozas, J., Björndahl, L., Menkveld, R., Lewis, S.E.M., Mortimer, D., Jorgensen, N., Barratt, C.L.R., Fernández, M.F., Castilla, J.A., 2013. Proposal of guidelines for the appraisal of SEMen QUAlity studies (SEMQUA). Hum. Reprod. 28, 10–21. https://doi.org/10.1093/humrep/des355.

Sharpe, RM, 2020. Androgens and the masculinization programming wondow: human-rodent differences. Biochem. Soc. Trans. 48, 1725.

Shi, M., Sekulovski, N., MacLean, J.A., Hayashi, K., 2018. Prenatal exposure to bisphenol A analogues on male reproductive functions in mice. Toxicol. Sci. 163, 620–631. https://doi.org/10.1093/toxsci/kfy061.

Shi, M., Whorton, A.E., Sekulovski, N., Maclean, J.A., Hayashi, K., 2019. Prenatal exposure to bisphenol A, E, and S induces transgenerational effects on male reproductive functions in mice. Toxicol. Sci. 172, 303–315. https://doi.org/10.1093/toxsci/kfz207.

Skakkebaek, N.E., Rajpert-De Meyts, E., Buck Louis, G.M., Toppari, J., Andersson, A.M., Eisenberg, M.L., Jensen, T.K., Jørgensen, N., Swan, S.H., Sapra, K.J., Ziebe, S., Priskorn, L., Juul, A., 2015. Male reproductive disorders and fertility trends: influences of environment and genetic susceptibility. Physiol. Rev. 96, 55–97. https://doi.org/10.1152/physrev.00017.2015.

Spörndly-Nees, E., Boberg, J., Ekstedt, E., Holm, L., Fakhrzadeh, A., Dunder, L., Kushnir, M.M., Lejonklou, M.H., Lind, P.M., 2018. Low-dose exposure to Bisphenol A during development has limited effects on male reproduction in midpubertal and aging Fischer 344 rats. Reprod. Toxicol. 81, 196–206. https://doi.org/10.1016/j.reprotox.2018.08.007.

Teuschler, L.K., Hertzberg, R.C., 1995. Current and future risk assessment guidelines, policy, and methods development for chemical mixtures. Toxicology 105, 137–144. https://doi.org/10.1016/0300-483X(95)03207-V.

Tyl, R.W., Myers, C.B., Marr, M.C., Sloan, C.S., Castillo, N.P., Veselica, M.M., Seely, J.C., Dimond, S.S., Miller, J.P. Van, Shiotsuka, R.N., Beyer, D., Hentges, S.G., Waechter, J. M., 2008. Two-Generation reproductive toxicity study of dietary bisphenol A in CD-1. Swiss) Mice 104, 362–384. https://doi.org/10.1093/toxsci/kfn084.

Tyl, R.W., Myers, C.B., Marr, M.C., Thomas, B.F., Keimowitz, A.R., Brine, D.R., Veselica, M.M., Fail, P.A., Chang, T.Y., Seely, J.C., Shiotsuka, R.N., 2002. Three-Generation reproductive toxicity study of dietary bisphenol A in CD. Sprague-Dawley Rats 146, 121–146.

Ullah, A., Pirzada, M., Jahan, S., Ullah, H., Razak, S., Rauf, N., 2019. Prenatal BPA and its analogs BPB , BPF , and BPS exposure and reproductive axis function in the male offspring of Sprague Dawley rats. https://doi.org/10.1177/0960327119862335.

Ullah, A., Pirzada, M., Jahan, S., Ullah, H., Turi, N., Ullah, W., Siddiqui, M.F., Zakria, M., Lodhi, K.Z., Khan, M.M., 2018. Impact of low-dose chronic exposure to bisphenol A and its analogue bisphenol B, bisphenol F and bisphenol S on hypothalamo-pituitary-testicular activities in adult rats: a focus on the possible hormonal mode of action. Food Chem. Toxicol. 121, 24–36. https://doi.org/10.1016/j.fct.2018.08.024.

Vilela, J, Hartmann, A, Silva, EF, Cardoso, T, Corcini, CD, Varela-Junior, AS, Martinez, PE, Colares, EP, 2014. Sperm impairments in adult vesper mice (*Calomys cauda*) caused by *in utero* exposure to bisphenol A. Andrologia 46, 971–978. https://doi.org/10.1111/and.12182.

Vitku, J., Heracek, J., Sosvorova, L., Hampl, R., Chlupacova, T., Hill, M., Sobotka, V., Bicikova, M., Starka, L., 2016. Associations of bisphenol A and polychlorinated biphenyls with spermatogenesis and steroidogenesis in two biological fluids from men attending an infertility clinic. Environ. Int. 89–90, 166–173. https://doi.org/10.1016/j.envint.2016.01.021.

Vom Saal, F.S., Cooke, P.S., Buchanan, D.L., Palanza, P., Thayer, K.A., Nagel, S.C., Parmigiani, S., Welshons, W.V., 1998. A physiologically based approach to the study of bisphenol a and other estrogenic chemicals on the size of reproductive organs, daily sperm production, and behavior. Toxicol. Ind. Health 14, 239–260. https://doi.org/10.1177/074823379801400115.

Whaley, P., Aiassa, E., Beausoleil, C., Beronius, A., Bilotta, G., Boobis, A., de Vries, R., Hanberg, A., Hoffmann, S., Hunt, N., Kwiatkowski, C.F., Lam, J., Lipworth, S., Martin, O.V., Randall, N., Rhomberg, L., Rooney, A.A., Schünemann, H.G., Wikoff, D., Wolffe, T., Halsall, C., 2020. Recommendations for the conduct of systematic reviews in toxicology and environmental health research (COSTER). Environ. Int. 143, 105926. https://doi.org/10.1016/j.envint.2020.105926.

Wang, H.F., Liu, M., Li, N., Luo, T., Zheng, L.P., Zeng, X.H., 2016. Bisphenol a impairs mature sperm functions by a CatSper-relevant mechanism. Toxicol. Sci. 152, 145–154. https://doi.org/10.1093/toxsci/kfw070.

WHO, 2010. https://apps.who.int/iris/handle/10665/44261. (Accessed 8 February 2021).

Yang, Y., Hong, Y., Chae, S.A., 2015. Reduction in semen quality after mixed exposure to bisphenol A and isobutylparaben in utero and during lactation periods 1–10. https://doi.org/10.1177/0960327115608927.

Yoshino, Hiroko, Ichihara, Toshio, Kawabe, Mayumi, Imai, Norio, Hagiwara, Akihiro, Asamoto, Makoto, Shirai, Tomoyuki, 2002. Lack of significant alteration in the prostate or testis of F344 rat offspring after transplacental or lactational exposure to bisphenol A. J. Toxicol. Sci. 27 (5), 433–439.