Taylor & Francis
Taylor & Francis Group

Critical Reviews in Toxicology

ISSN: (Print) (Online) Journal homepage: https://www.tandfonline.com/loi/itxc20

Toward a science-based testing strategy to identify maternal thyroid hormone imbalance and neurodevelopmental effects in the progeny – part I: which parameters from human studies are most relevant for toxicological assessments?

Ursula G. Sauer , Alex Asiimwe , Philip A. Botham , Alex Charlton , Nina Hallmark , Sylvia Jacobi , Sue Marty , Stephanie Melching-Kollmuss , Joana A. Palha , Volker Strauss , Bennard van Ravenzwaay & Gerard Swaen

To cite this article: Ursula G. Sauer , Alex Asiimwe , Philip A. Botham , Alex Charlton , Nina Hallmark , Sylvia Jacobi , Sue Marty , Stephanie Melching-Kollmuss , Joana A. Palha , Volker Strauss , Bennard van Ravenzwaay & Gerard Swaen (2020): Toward a science-based testing strategy to identify maternal thyroid hormone imbalance and neurodevelopmental effects in the progeny – part I: which parameters from human studies are most relevant for toxicological assessments?, Critical Reviews in Toxicology

To link to this article: https://doi.org/10.1080/10408444.2020.1839380

CRITICAL REVIEWS IN TOXICOLOGY
https://doi.org/10.1080/10408444.2020.1839380

Taylor & Francis
Taylor & Francis Group

REVIEW ARTICLE

OPEN ACCESS Check for updates

Toward a science-based testing strategy to identify maternal thyroid hormone imbalance and neurodevelopmental effects in the progeny – part I: which parameters from human studies are most relevant for toxicological assessments?

Ursula G. Sauer[a], Alex Asiimwe[b], Philip A. Botham[c], Alex Charlton[c], Nina Hallmark[d], Sylvia Jacobi[e], Sue Marty[f], Stephanie Melching-Kollmuss[g], Joana A. Palha[h,i,j], Volker Strauss[g], Bennard van Ravenzwaay[g] and Gerard Swaen[k]

[a]Scientific Consultancy – Animal Welfare, Neubiberg, Germany; [b]Bayer AG, Berlin, Germany; [c]Syngenta, Jealott's Hill, Bracknell, UK; [d]Bayer AG, Monheim, Germany; [e]Albemarle Europe SRL, Louvain-la-Neuve, Belgium; [f]The Dow Chemical Company, Midland, MI, USA; [g]BASF SE, Ludwigshafen, Germany; [h]Life and Health Sciences Research Institute (ICVS), School of Medicine, University of Minho, Braga, Portugal; [i]ICVS/3B's, PT Government Associate Laboratory, Guimaraes, Portugal; [j]Clinical Academic Center, Braga, Portugal; [k]Department of Complex Genetics, Maastricht University, Maastricht, The Netherlands

ABSTRACT

The 2018 European Food Safety Authority/European Chemicals Agency Guidance on the Identification of Endocrine Disruptors lacks clarity on how the presence or absence of substance-induced maternal thyroid hormone imbalance, or the potential for subsequent deleterious consequences in child neurodevelopment, should be established by toxicological assessments. To address these uncertainties, this narrative review evaluates human evidence on how altered maternal thyroid function may be associated with child neurodevelopmental outcomes; and seeks to identify parameters in human studies that appear most relevant for toxicological assessments. Serum levels of free thyroxine (fT4) and thyroid stimulating hormone (TSH) are most frequently measured when assessing thyroid function in pregnant women, whereas a broad spectrum of neurodevelopmental parameters is used to evaluate child neurodevelopment. The human data confirms an association between altered maternal serum fT4 and/or TSH and increased risk for child neurodevelopmental impairment. Quantitative boundaries of effects indicative of increased risks need to be established. Moreover, it is unknown if altered serum levels of total T4, free or total triiodothyronine, or parameters unrelated to serum thyroid hormones might be more relevant indicators of such effects. None of the human studies established a link between substance-mediated liver enzyme induction and increased serum thyroid hormone clearance, let alone further to child neurodevelopmental impairment. This review identifies research needs to contribute to the development of toxicity testing strategies, to reliably predict whether substances have the potential to impair child neurodevelopment via maternal thyroid hormone imbalance.

ARTICLE HISTORY
Received 5 June 2020
Revised 22 September 2020
Accepted 16 October 2020

KEYWORDS
Free thyroxine (fT4); thyroid stimulating hormone (TSH); thyroid disruption; developmental neurotoxicity; uridine diphosphate glucuronyltransferase (UGT); adverse outcome pathway (AOP); brain development; pregnancy; European Union guidelines; endocrine disruptors

Table of contents

Introduction ... 2
Objective 1: evaluate human evidence for how maternal thyroid hormone levels and child neurodevelopmental effects may be linked ... 4
Objective 2: evaluate human evidence for impact of liver enzyme induction on maternal serum T4 (and TSH) levels and, ultimately, child neurodevelopmental impairment ... 4
Overarching goal: identify parameters in human studies, which are most relevant for toxicological assessments ... 4
Methodology: literature searches and selection of relevant literature ... 5
Objective 1: human evidence for how maternal thyroid hormone levels and child neurodevelopment may be linked ... 5
Objective 2: human evidence for impact of liver enzyme induction on maternal serum T4 (and TSH) levels and, ultimately, child neurodevelopment ... 5
Summary of the human evidence reviewed ... 6
Human evidence for how maternal TSH and fT4 levels and child neurodevelopment may be linked ... 6
Evidence from the systematic reviews and meta-analyses ... 6
Evidence from the randomised clinical trial ... 7
Evidence from the observational studies ... 7

CONTACT Bennard van Ravenzwaay bennard.ravenzwaay@basf.com BASF SE, RB/T, Z470, Carl-Bosch-Strasse 38, Ludwigshafen, 67056, Germany

Human evidence for impact of liver enzyme induction on serum T4 (and TSH) levels ... 9
Evidence for impact of maternal exposure to liver enzyme-inducing substances on child neurodevelopment ... 11
Evidence for impact of exposure to substances, which induce liver enzymes, on serum T4 (and TSH) levels ... 11
Reviews addressing the impact of exposure to liver enzyme-inducing substances on serum T4 (and TSH) levels ... 12
Discussion ... 12
Objective 1: human evidence for how maternal thyroid hormone levels and child neurodevelopment may be linked ... 12
Do internationally accepted normal values (ranges) for serum TSH or thyroid hormone levels in pregnant women exist? If not, are they needed e.g. for the interpretation of human studies? ... 12
Is it possible to identify thresholds for altered serum TSH or fT4 levels in pregnant women indicating an increased risk for child neurodevelopmental impairment? ... 14
Is there evidence for a correlation between altered serum TSH or thyroid hormone levels in pregnant women without a history of thyroid disease and increased risk for child neurodevelopmental impairment? ... 15
Is there evidence that T4 supplementation in healthy pregnant women with high TSH and/or low T4 serum levels decreases the risk for neurodevelopmental impairment in their children? ... 15
Can the most sensitive parameter(s) addressing child neurodevelopmental outcomes be identified? ... 16
Objective 2: human evidence for impact of liver enzyme induction on maternal serum T4 (and TSH) levels and, ultimately, child neurodevelopmental impairment 16
Is it physiologically plausible that liver enzyme induction has an impact on maternal serum thyroid hormone levels and, ultimately, child neurodevelopment? If so, to which extent would such an effect be compensated for by the TSH feedback system? ... 16
Summary of the findings and conclusions ... 17
Which serum parameters appear most relevant to identify maternal thyroid dysfunction? ... 17
Summary of the findings ... 17
Conclusion ... 17
Which further (non-serum) parameters might be relevant to identify maternal thyroid dysfunction? ... 18
Summary of the findings ... 18
Conclusion ... 18
Which parameters appear most relevant to identify child neurodevelopmental impairment? ... 19
Summary of the findings ... 19
Conclusion ... 19
Which liver function parameters appear relevant to identify substance-mediated liver enzyme induction leading to increased serum thyroid hormone clearance? ... 19
Summary of the findings ... 19
Conclusion ... 19
Final conclusion and next steps ... 19
Acknowledgements ... 20
Declaration of interest ... 20
ORCID ... 21
References ... 21

Introduction

Toxicological assessments serve the goal of protecting humans and the environment from unwanted substance-induced effects. Animal toxicity studies, *in silico* modeling, and *in vitro* assays are performed to predict effects that could occur in humans if they were exposed to the respective substances (Wilks et al. 2015). Recently, concerns have been expressed that chemical-induced alterations in serum thyroid hormone levels in pregnant women may impair child neurodevelopment, e.g. "How chemicals can affect the health of developing children" (ECHA 2018). These concerns raise the question whether available toxicological test methods are reliable in predicting such effects. Any prevailing knowledge gaps must be identified so that they can be addressed to ensure the safe use of chemicals. The present review aims to contribute to this goal by determining which specific effects in pregnant women and their children need to be reflected in toxicological assessments to enable scientifically sound safety assessments.

The thyroid gland is an endocrine organ present in all vertebrates. Thyroxine (T4) is the main thyroid hormone synthesised and secreted by the thyroid, whereas triiodothyronine (T3), the biologically active hormone, is mostly produced by deiodination of T4 in peripheral tissues. A major role of the thyroid hormones is to regulate metabolism, e.g. during growth and reproduction. In developing offspring, rodent data (with supporting evidence in humans) indicate that thyroid hormones play a role in neuronal migration, cellular differentiation (e.g. of neurons) and glial myelination. The thyroid gland is controlled by the pituitary [through secretion of thyroid stimulating hormone (TSH)] which, in turn, is regulated by the thyrotropin releasing hormone secreted by the hypothalamus (Dickhoff and Darling 1983; DeGroot and Jameson 2001). Thyroid hormones are highly hydrophobic and therefore generally bound to serum binding proteins when circulating in the bloodstream, whereas only a minor fraction (<1%) remains as free hormones (e.g. free T4 (fT4) and free T3). It is the free hormone fraction that is sensed by the tissues, triggering the homeostatic regulatory mechanisms (Stockigt 2001).

The concern that chemical-induced alterations in serum thyroid hormone levels in pregnant women may impair child

neurodevelopment is reflected in the European Food Safety Authority (EFSA) and European Chemicals Agency (ECHA) "Guidance for the identification of endocrine disruptors in the context of Regulations (EU) No 528/2012 and (EC) No 1107/2009" that was published in June 2018 (EFSA and ECHA 2018). Therein, Appendix A "Additional considerations for how to assess the potential for thyroid disruption for human health" singles out concerns for thyroid disruption.

Various statements in this Appendix A suggest that (1) there is unambiguous evidence that low maternal serum T3 and/or T4 levels *per se* result in child neurodevelopmental impairment regardless of whether the histology and/or function of the thyroid gland is also affected, and (2) that this can clearly be identified in rodent studies. The report of a European Commission Workshop on Thyroid Disruption (European Commission 2017a) is used as the main reference in Appendix A to substantiate this view. Furthermore, Appendix A proposes that the human relevance of the effects observed in animals "could be further investigated," specifying that this investigation should address whether the effects result from the induction of metabolic liver enzymes, such as T4-uridine diphosphate glucuronyltransferase (T4-UGT), leading to increased serum T4 clearance (EFSA and ECHA 2018).

A specific subsection of Appendix A is dedicated to the consequences of liver enzyme induction on thyroid hormone metabolism. This contrasts with the brief mention to other molecular initiating events leading to thyroid disruption, such as the inhibition of sodium-iodide symporter or the inhibition of thyroid peroxidase (TPO) activity, which both influence thyroid hormone synthesis (Noyes et al. 2019). Similarly, reduced serum T4 levels are the focus of Appendix A, whereas high serum T4 levels are not explicitly mentioned.

Appendix A describes a testing scheme to determine serum thyroid hormone levels and liver enzyme activities. However, it does not indicate how the various parameters should be measured or how the data should be evaluated within a weight-of-evidence approach to reach a conclusion on whether, or not, a substance fulfills the criteria of an endocrine disruptor, as implemented for biocides and plant protection products in the European Commission (2017b, 2018) Regulations.

Hence, based upon EFSA and ECHA (2018) Guidance, it is currently unclear how the presence or absence of substance-induced maternal thyroid disruption (with or without liver enzyme induction as the initiating event), and potential deleterious consequences in child neurodevelopment, should be established by toxicological assessments.

To address this scientific and regulatory uncertainty, in July 2018, the European Centre for Ecotoxicology and Toxicology of Chemicals (ECETOC) convened a Special T4 Task Force. It is the goal of this Task Force to contribute to the development of a science-based testing strategy to identify if a substance has the potential to elicit maternal thyroid hormone imbalances and subsequently neurodevelopmental effects in the child. In pursuing this goal, the ECETOC Special T4 Task Force is preparing a series of reviews, of which the present review constitutes the first. These reviews shall include:

- An evaluation of the human evidence on how lower serum levels of maternal thyroid hormone affect child neurodevelopment including cases where substance-mediated liver enzyme induction leads to increased thyroid hormone clearance (presented herein).
- An evaluation of how molecular initiating events and key events of thyroid-related adverse outcome pathways (AOPs) that include neurodevelopmental adverse outcomes are being addressed in *in vivo* and/or *in vitro* toxicological assessments. This evaluation shall include quantitative aspects of key event relationships and possible species differences in key events and key event relationships (Noyes et al. 2019).
- An evaluation of the evidence from available rodent studies showing how specific xenobiotic substances affect the maternal thyroid hormone system, and further, under which conditions such effects lead to (which types of) neurodevelopmental effects in the pups, and how these are predictive of effects in humans. One of the uncertainties in associating thyroid hormone imbalances with adverse neurodevelopmental consequences is that, in rats, the methods to measure neurodevelopmental outcomes are much less refined than the methods used in children.

Finally, all collated evidence shall be consolidated to propose a science-based tiered testing strategy to identify if a substance has the potential to elicit imbalances of the maternal thyroid hormone system, if such imbalances might further lead to neurodevelopmental impairment of the progeny, and, if so, beyond which boundaries of altered thyroid hormone levels such effects may be expected. Research needs to enhance the applicability of the testing strategy shall be identified.

The present first review addressing the human evidence on how lower maternal thyroid hormone levels affect child neurodevelopment was aligned with the topics of Appendix A of the EFSA and ECHA (2018) Guidance. With respect to maternal serum parameters, hypothyroidism (low fT4 with concordant high TSH) and hypothyroxinaemia (isolated low fT4) were addressed. Importantly, focus is not on severe states of congenital hypothyroidism caused by e.g. pronounced maternal iodine deficiency leading to cretinism (Boyages and Halpern 1993), but rather on more subtle effects on maternal thyroid function.

Hyperthyroidism (low TSH with concordant high fT4) and hyperthyroxinaemia (isolated high fT4) were excluded from the scope of this review since they are not specifically addressed in Appendix A. Nevertheless, hyperthyroidism and hyperthyroxinaemia are also relevant from a clinical perspective (Korevaar et al. 2016a; Jansen et al. 2019). Similarly, with respect to initiating events, this review considers how substance-mediated liver enzyme induction affects serum thyroid hormone levels in humans. By contrast, other important initiating events that affect serum thyroid hormone levels and, ultimately, child neurodevelopmental outcomes, such as TPO

activity or iodine deficiency, are not addressed here for not being considered in Appendix A. Nonetheless, and as mentioned above, the iodine status plays a critical role in ensuring physiological thyroid function.

Further, this review focuses on the first trimester of pregnancy when the embryo is still fully dependent upon maternal thyroid hormone supply (Fisher and Klein 1981; Thorpe-Beeston et al. 1991; Choksi et al. 2003). During this critical period of development, reduced maternal serum T4 levels are most likely to have the greatest impact on child neurodevelopment (Costeira et al. 2011; Murcia et al. 2011, 2018; Levie et al. 2019a).

In exploring how (liver enzyme induction-mediated) thyroid hormone imbalance in pregnant women affects child neurodevelopment, the ECETOC Special T4 Task Force first produced a preliminary review that was then further discussed with invited experts, including clinicians, representatives from authorities, regulatory toxicologists and veterinary pathologists from Europe and North America, at an Extended Task Force Meeting that took place on 21 and 22 November 2019 in Ludwigshafen, Germany. The present article reflects the combined outcome of both activities.

Importantly, a narrative review (see Supplementary Information SI-1 for definition) of the published literature was conducted, and not a systematic review. Its goal was to exemplarily identify, evaluate and synthesise relevant human data, from which a reliable overview can be provided on how altered maternal thyroid hormone levels lead to child neurodevelopmental impairment, including the potential for liver enzyme induction to affect circulating thyroid hormone levels. The human studies covered include epidemiological studies, i.e. observational cohort, case–cohort and nested case–control studies, as well as a randomised clinical trial (see Supplementary Information SI-1 for definitions). The underlying hypothesis is that information from such human studies is useful to identify parameters that are relevant for, and hence should be reflected in, toxicological assessments.

Epidemiology is the discipline addressing the distribution of disease among human populations thereby aiming at identifying its causes and contributory factors (Hajat 2011). The scientific methodologies available in epidemiology are fundamentally different from those available for toxicological assessments. Toxicologists assessing chemical-induced effects can apply experimental study designs and testing strategies using a broad spectrum of test methods, many of which are standardised and adopted as internationally agreed formal guidance. By contrast, epidemiologists, when evaluating human health effects potentially caused by chemical exposure, are nearly always limited to non-experimental, observational study designs (WHO IPCS 2004), and the types of health parameter measurements that can be made are much more restricted than e.g. in regulatory toxicology.

With these limitations in mind, the present narrative review pursued two objectives, the outcomes of which were then further evaluated in view of the overarching goal to identify parameters from human studies, which appear most relevant for toxicological assessments.

Objective 1: evaluate human evidence for how maternal thyroid hormone levels and child neurodevelopmental effects may be linked

(1) Do internationally accepted normal values (ranges) for serum TSH or thyroid hormone levels in pregnant women exist? If not, are they needed for the interpretation of human studies (e.g. to enhance between-study comparability)?
(2) Is it possible to identify thresholds for high TSH or low thyroid hormone levels in the serum of pregnant women indicating an increased risk for child neurodevelopmental impairment? Throughout the text, the term "threshold" is used in the meaning of "quantitative boundaries for high TSH and/or low thyroid hormone levels in the serum of pregnant women beyond which an increased risk for child neurodevelopmental impairment may be expected." While high thyroid hormone and low TSH are not considered in the present review, the term "quantitative boundary" reflects the expectation that there is an optimum range of maternal serum fT4 and/or thyroid hormone within which there is no such increased risk. Further, the term considers that there is never no such risk, also because thyroid hormone imbalances are not the only potential causes for neurodevelopmental impairment.
(3) Is there evidence for a correlation between (i) high TSH and/or low thyroid hormone serum levels in pregnant women without a history of thyroid disease and (ii) child neurodevelopmental impairment?
(4) Is there evidence that T4 supplementation in pregnant women with high TSH and/or low T4 serum levels decreases the risk for neurodevelopmental impairment in their children?
(5) Can the most sensitive parameter(s) addressing child neurodevelopmental outcomes be identified?

Objective 2: evaluate human evidence for impact of liver enzyme induction on maternal serum T4 (and TSH) levels and, ultimately, child neurodevelopmental impairment

(6) Is it physiologically plausible that substance-mediated liver enzyme induction in humans has an impact on maternal serum thyroid hormone levels and, ultimately, child neurodevelopment? If so, to which extent would such effects be compensated for by the TSH feedback system?

Overarching goal: identify parameters in human studies, which are most relevant for toxicological assessments

Based upon the evidence collated and evaluated in pursuing Objectives 1 and 2, the final section of this review refers back to the overarching question raised in its title *Which parameters from human studies are most relevant for toxicological assessments?* Scientific evidence is needed to decide on how to assess if chemicals have the potential to disrupt thyroid function with deleterious consequences to the health of mothers and their progeny. Specifically, the human evidence

collated in pursuing the above-mentioned six questions is utilised to identify:

- Serum and non-serum parameters that appear most relevant to identify maternal thyroid disruption;
- Parameters that appear most relevant to identify child neurodevelopmental outcomes; and
- Liver function parameters that appear most relevant to identify substance-mediated liver enzyme induction leading to increased serum thyroid hormone clearance.

This evaluation of the human evidence shall contribute to the development of the planned tiered testing strategy to identify if a substance has the potential to elicit imbalances of the (maternal) thyroid hormone system, and if this might subsequently result in impaired child neurodevelopment, as well as if effects observed in rodents are relevant for humans. While this work of the ECETOC Special T4 Task Force is motivated by the knowledge gaps identified in Appendix A of EFSA and ECHA (2018) Endocrine Disruptor Guidance, the testing strategy shall not only be applicable for substances regulated under the EU Biocidal Products and Plant Protection Products Regulations, as covered by EFSA and ECHA (2018). Rather, by founding the testing strategy on the state-of-the-art science and technology (which shall be established in the planned series of reviews of which the present one constitutes the first), it should be employable irrespective of the responsible jurisdiction, or applicable legislation.

Methodology: literature searches and selection of relevant literature

The methodology applied to conduct the literature searches and to identify relevant literature was selected in view of the goal to conduct a narrative review (Supplementary Information SI-1) and was not intended to comply with rules for systematic reviews. Sources of information include literature searches conducted in the US National Institutes of Health National Library of Medicine database PubMed (https://ncbi.nlm.nih.gov/pubmed/ [accessed 2020 September]). Further, the references listed in relevant articles were checked for additional literature of interest ("snowballing"). In this regard, the Discussion Section also considered references from the two documents that motivated the present work, i.e. Appendix A of EFSA and ECHA (2018) and the European Commission (2017a) Thyroid Disruption Workshop Report.

All literature retrieved that fulfilled the inclusion criteria defined empirically *a priori* in pursuing Objectives 1 and 2 was included (see below for details). While the human evidence was evaluated from a toxicological perspective (i.e. in view of identifying biological processes and biomarkers (parameters) that should be reflected in toxicity studies), the inclusion criteria relate to "Population, Intervention, Comparison, Outcome and Study type," as recommended for epidemiological studies (Tacconelli 2010; Methley et al. 2014).

Objective 1: human evidence for how maternal thyroid hormone levels and child neurodevelopment may be linked

The following search query was used:

(thyroid OR thyrox OR T4) AND ("meta-analysis" OR "systematic review" OR cohort OR trial) AND (pregnan* OR fetus OR maternal) AND (neurodevelop* OR cognitive OR IQ OR "brain develop*")*

This search query yielded 217 retrievals on 6 September 2018, and 42 further documents on 28 March 2020.

Articles were pre-selected based upon an evaluation of title and abstract, if they had been published between 2008 and March 2020, or earlier if retrieved by snowballing, and fulfilled the following criteria that were defined empirically *a priori*:

- Population: Pregnant women and their offspring.
- Intervention: Maternal blood sampling during the first trimester of pregnancy (median gestational week (GW) $\leq$ 13.3) as critical window of susceptibility when the embryo is still fully dependent upon maternal thyroid hormone supply (Fisher and Klein 1981; Thorpe-Beeston et al. 1991; Choksi et al. 2003) and assessment of maternal serum T4 and/or TSH levels.
- Comparison: Mothers with study-specific normal T4 and/or TSH ranges versus mothers with study-specific high TSH and/or low T4 (hypothyroidism and/or hypothyroxinaemia).
- Outcome: Child neurodevelopmental impairment; any type of parameter.
- Study type: Systematic reviews, meta-analyses, any type of observational human study, randomised clinical trials; sample sizes of (generally) $\geq$ 1,000 mother–child pairs.

These selection criteria served to ensure that all human studies included in the evaluation were of sufficient, comparable quality. To ensure comprehensiveness in pursuing the aim to identify parameters in human studies that should be reflected in toxicological assessments, all studies retrieved that met the criteria were included in the evaluation.

After full-text evaluation, one meta-analysis (Spencer et al. 2015) retrieved was excluded as out of scope since it aimed at evaluating the overall benefit of thyroid screening programmes during pregnancy, i.e. maternal fT4 and TSH were primarily related to pregnancy outcomes, such as risks for pre-eclampsia, preterm birth, or miscarriage, but not to child neurodevelopment. Observational human studies were excluded as out of scope if maternal blood sampling had been undertaken after the first trimester of pregnancy (Haddow et al. 1999; Li et al. 2010; Chevrier et al. 2011) or if only neonatal, but not maternal blood samples had been taken.

Objective 2: human evidence for impact of liver enzyme induction on maternal serum T4 (and TSH) levels and, ultimately, child neurodevelopment

The focus of this search query was on human studies addressing whether and how exposure to antiepileptic drugs mediates liver enzyme induction and is then followed by

increased T4 clearance. Antiepileptic drugs were selected as an exemplary substance group since they are widely applied in humans, so that a solid database was expected, and since a number of antiepileptic drugs have reported hepatic and thyroid side-effects (Hamed 2015). Due to the overall paucity of relevant data (see below), this database was supplemented by human studies addressing hepatic and thyroid effects of non-pharmaceutical substances.

PubMed search queries were conducted as indicated below. If a search query yielded 150 retrievals or more, it was narrowed down by adding further search terms and/or by applying the automatic PubMed filter "human." Since it was not the goal to conduct a systematic literature review, this automatic filtering appeared justifiable.

- *(thyroid OR thyrox*) AND ("antiepileptic" OR "phenobarbital")*: 435 retrievals
  - with filter "human": 206 retrievals (both: on 10 September 2018); therefore, the following further restriction was applied:
- *(thyroid OR thyrox*) AND ("antiepileptic" OR "phenobarbital") AND (liver OR hepatic OR UGT OR UDPGT OR metabol*)*: 295 retrievals
  - with filter "human": **116 retrievals** (both: on 10 September 2018)
- *"UGT (thyroid* OR thyrox*)"*: **83 retrievals** and *"(UDPGT NOT UGT) (thyroid* OR thyrox*)"*: **36 retrievals** (both: on 28 August 2018)
- *"UGT AND (neurodevelopment OR cognitive) (hepatic OR liver)"*: **1 retrieval** on 28 August 2018, i.e. Builee and Hatherill (2004)
- *"liver enzyme" AND (neurodevelopment OR cognitive) (hepatic OR liver):* **12 retrievals** on 28 August 2018; none of which were relevant for Objective 2. If retrievals presented human studies, these generally addressed alcohol abuse or schizophrenia unrelated to substance exposure.

If search strings yielded less than 150 retrievals (indicated above in bold), their titles and abstracts were evaluated to identify potentially relevant articles that were then submitted to full text evaluation. The automatic PubMed filter "human" also captured e.g. genomics or *in vitro* studies using human genes, cells, or tissues. Such articles were excluded as out of scope.

Articles met the inclusion criteria if they were published between 1980 and 2018 (approx. four decades to account for the overall paucity of relevant information), or earlier if retrieved by snowballing, and fulfilled the following criteria that were defined empirically *a priori*:

- Population:
  - Initial search criterion: Pregnant mothers and their children;
  - Expanded search criteria (to account for the paucity of relevant literature retrieved applying the initial search criterion): All humans regardless of gender or age (exception: breast-fed infants).
- Intervention: Exposure to liver enzyme-inducing substances:
  - Initial search criterion: Indirect human exposure to antiepileptic drugs, i.e. drug-mediated induction of maternal liver UGT leading to maternal thyroid hormone imbalances and, possibly, child neurodevelopmental impairment;
  - Expanded search criteria: Direct human exposure to liver enzyme inducing-substances with a focus on, but not restricted to, antiepileptic drugs; relevant population: exposed individuals regardless of age or gender; exception: exposure via lactation.
- Comparison: Exposed humans versus non-exposed humans.
- Outcome: (1) Liver enzyme induction; (2) increased serum T4 clearance (and secondarily altered TSH levels); and (3) in case of indirect (*in utero*) exposure, neurodevelopmental outcome (any type of parameter).
- Study type: Reviews, systematic reviews, meta-analyses, observational studies.

Summary of the human evidence reviewed

Human evidence for how maternal TSH and fT4 levels and child neurodevelopment may be linked

Below, the human evidence reviewed in pursuing Objective 1 is summarised. All of the evaluated articles used TSH and fT4 as measures of maternal thyroid function. The presence of TPO antibody positivity was determined in some studies, total T4 (tT4) in a few, and in one study also total T3. None of the articles included specific evaluations of the thyroid gland, e.g. ultrasound examination or histopathological analysis of biopsies.

Evidence from the systematic reviews and meta-analyses

Two systematic reviews (Fetene et al. 2017; Drover et al. 2019) and five meta-analyses (Fan and Wu 2016; Wang et al. 2016; Levie et al. 2018, 2019b; Thompson et al. 2018) were included in the evaluation. Generally, these publications suggested a link between high TSH and/or low T4 levels in the serum of pregnant mothers and child neurodevelopmental impairment.

The only exception is the meta-analysis by Levie et al. (2019b), which evaluated data from three cohort studies together, i.e. (1) the *Infancia y Medio Ambiente* (Spain); (2) the Generation R (The Netherlands); (3) the Avon Longitudinal Study of Parents and Children (United Kingdom); for further details on these three cohort studies, see below Section *Evidence from the observational studies*. Levie and coworkers found no clear evidence for an association between maternal TSH and fT4 up until GW 18 and child attention deficit hyperactivity disorder (Levie et al. 2019b). By contrast, the systematic review by Drover et al. (2019), that included 28 studies (5 of which addressing neonatal thyroid hormone levels), reported moderate evidence for such an association.

Overall, none of the systematic reviews and meta-analyses revealed quantitative boundaries for high TSH and/or low T4 levels in the serum of pregnant mothers indicative of increased risks for child neurodevelopmental impairment.

Such boundaries could not be determined for various reasons. On the one hand, the systematic reviews and meta-analyses, as well as the underlying epidemiological studies, were not designed to identify quantitative boundaries. Therefore, the respective data were not collected and processed in view of meeting such a goal. Further, study design of the individual epidemiological studies varied greatly, as was also generally denoted in the systematic reviews and meta-analyses.

Evidence from the randomised clinical trial

The randomised clinical trial was conducted in the UK (Lazarus et al. 2012; Hales et al. 2018) and Italy (Lazarus et al. 2012) and assessed whether antenatal treatment for hypothyroidism was beneficial to child cognitive function (intelligence quotient, IQ) at the age of 3 years (Lazarus et al. 2012) and 9.5 years (Hales et al. 2018). "Suboptimal gestational thyroid function" requiring treatment was identified by either low maternal serum fT4 (<8.4 pmol/L corresponding to <2.5th percentile), or elevated TSH (>3.65 mU/L or >97.5th percentile), or a combination of both. Treatment for hypothyroidism (50 μg levothyroxine/day to achieve a target serum TSH level of 0.1–1.0 mU/L) was started in the positive women at GW 13 and 3 days (median). The IQ of children of mothers with normal gestational thyroid function did not differ from that of children of mothers with either treated or untreated suboptimal gestational thyroid function. Both Lazarus et al. (2012) and Hales et al. (2018) concluded that antenatal treatment for suboptimal gestational function without overt hypothyroidism did not result in improved cognitive function while cautioning that the treatment might have been initiated too late in pregnancy to positively affect this (further addressed in the Discussion Section).

Evidence from the observational studies

The prospective observational studies considered in this narrative review used data from nine different mother–child cohorts:

1. The Generation R cohort from Rotterdam, Netherlands (Henrichs et al. 2010; Ghassabian et al. 2011; Korevaar et al. 2016a; Jansen et al. 2019);
2. An earlier Dutch cohort study using 220 mother–child pairs (Pop et al. 1999);
3. The Spanish Environment and Childhood (*Infancia y Medio Ambiente*) project cohort (Julvez et al. 2013);
4. The Millennium cohort study from Scotland, United Kingdom (Williams et al. 2013);
5. The Northern Finland birth cohort 1986 (Päkkilä et al. 2015);
6. The Avon Longitudinal Study of Parents and Children cohort, United Kingdom (Fetene et al. 2018, 2020; Nelson et al. 2018);
7. A Danish case–cohort study investigating child neurodevelopmental disorders (Andersen et al. 2018);
8. A nested case–control study embedded in the Child Health and Development Study birth cohort from California, USA, investigating child bipolar disorder (Spann et al. 2020);
9. The Project Viva cohort from Massachusetts, USA (Oken et al. 2009; Lain et al. 2020).

In these studies, maternal blood samples had been taken between GW 9 and 13. Neurological assessments in the offspring were undertaken when the children were 10 months to 16 years old. Overall, the study designs varied greatly. Differences relate to the criteria to select mother–child pairs, to the reference levels and ranges applied to determine (different levels of) maternal hypothyroidism and/or hypothyroxinaemia (Table 1), and to the neurodevelopmental parameters applied to assess child neurodevelopmental outcomes (see Supplementary Information SI-2 for further details on cohort eligibility criteria, confounders, and statistical analysis procedures).

With respect to reference levels of maternal thyroid function, Pop et al. (1999) and Henrichs et al. (2010) focused on mothers with "mild hypothyroxinaemia" and "severe hypothyroxinaemia" and defined these as fT4 levels below the 10th and 5th percentiles of the respective cohort-specific normal ranges, respectively. In Pop et al. (1999), these percentiles correspond to <9.8 and <10.4 pmol/L, respectively, and in Henrichs et al. (2010) to ≤10.96 and ≤11.76 pmol/L, respectively. Ghassabian et al. (2011) and Spann et al. (2020) defined "hypothyroxinaemia" by fT4 reference levels below the respective 10th percentiles (corresponding to <11.76 pmol/L in the study by Ghassabian et al. and ≤0.86 ng/dL in the study by Spann et al.) and Julvez et al. (2013) by a fT4 level below the 5th percentile (corresponding to <8.39 pmol/L). Andersen et al. (2018) and Nelson et al. (2018) categorised mothers into sub-cohorts exhibiting subclinical hypothyroidism, overt hypothyroidism or hypothyroxinaemia and sought to identify statistically significant differences in child neurodevelopmental outcomes between these sub-cohorts. Päkkilä et al. (2015) did not refer to percentiles, but defined fT4 levels of <11.4 pmol/L and TSH levels of >3.1 mU/L as outside the normal range. Korevaar et al. (2016a) aimed at investigating "the shape of the association between TSH or fT4 concentrations and child IQ or brain MRI" (Table 1).

The human studies also varied greatly with respect to the neurodevelopmental parameters and statistical analysis methodologies. Neurodevelopmental assessments included psychomotor and mental development, cognitive function (i.e. IQ measurements), expressive vocabulary and educational attainment, brain magnetic resonance imaging (MRI) and diagnosis of autism, attention deficit hyperactivity disorder or bipolar disorder (see Supplementary Information SI-2 for further details).

While the assessments were mostly undertaken by trained persons, in some studies they were undertaken by the mothers or both parents (Henrichs et al. 2010; Ghassabian et al. 2011; Päkkilä et al. 2015), or by the children themselves (Päkkilä et al. 2015).

In spite of the differences between the neurodevelopmental parameters, the majority of studies reports an association between low maternal serum fT4 (and

Table 1. Maternal thyroid hormone reference levels and/or ranges indicated in the evaluated studies.

		Median, reference range		Hypothyroidism		Hypothyroxinaemia
Publication	GW and GW range	TSH [mU/L]	fT4 [pmol/L; or noted]	Overt	Subclinical	fT4: pmol/L
Meta-analyses, systemic review, workshop report						
Fan and Wu (2016)	Listed for individual cohorts	Distinguishes between subclinical hypothyroidism and hypothyroxinaemia, but does not provide definitions for reference levels or normal ranges (also not from the individual studies)				
Wang et al. (2016)	Not addressed	Lists the definitions for hypothyroidism and hypothyroxinaemia provided in the individual studies (but not the respective normal references values)				
European Commission (2017a)	Not applicable	4–10	Not addressed	TSH > 10 mU/L and T4 < reference range	Mild: Normal TSH; fT4 < 10 pctl Severe: TSH > 10 mU/L OR fT4 < 5 pctl	
Fetene et al. (2017)	Indicates trimester	Refers to hypothyroidism, subclinical hypothyroidism, mild and severe hypothyroxinaemia but without defining these terms (or mentioning the respective definitions from the analyzed studies)				
Thompson et al. (2018)	Not addressed	"high" and "low" TSH or fT4 / tT4 not defined		"high" TSH and "low" fT4 or tT4	"high" TSH and normal fT4 or tT4	Normal TSH and "low" fT4 or tT4
Levie et al. (2018)	Lists means	2.5–97.5 pctl	2.5–97.5 pctl	Not defined	TSH > 97.5 pctl fT4 normal	<2.5 pctl
Cohort, case–cohort and nested case–control studies						
Pop et al. (1999)	sampling at GW 12*	Included, but not recorded	13.2	Not defined	Not defined	<9.8 = lowest 5 pctl <10.4 = lowest 10 pctl
Oken et al. (2009)	10.2*	0.35–5.5	tT4: 4.5–10.9 µg/dL	Dichotomised as <2.5 or ≥2.5	Not defined	Not defined
Henrichs et al. (2010)	13.3 ± 1.7†	0.03–2.5	11–25	Not defined	Not defined	Mild: <11.76 = <10 pctl Severe: ≤10.96 = <5 pctl
Ghassabian et al. (2011)	13.3 ± 1.7†	0.03–2.5	11–25‡	Not defined	Not defined	<11.76 = <10 pctl
Julvez et al. (2013)	13 8–20†	0.50–2.59 10–90 pctl	8.9–12.3 10–90 pctl	Not defined	Not defined	<8.39 = <5 pctl
Williams et al. (2013)	10*	1.7 (1.2–3.2)	15.0 (SD: 1.7)¶	Not defined	Mild: TSH ≥ 2.5 mU/L	Not defined
Päkkilä et al. (2015)	10.7 ± 2.8*	0.07–3.1§	11.4–22.4	TSH > 3.1 fT4 < 11.4	TSH > 3.1 fT4 normal	TSH normal fT4 < 11.4
Korevaar et al. (2016a)	13.2; 95% range: 9.8–17.5*	1.35 (0.05–4.96)	14.9 (10.2–22.4)	TSH > 97.5 pctl fT4 < 2.5 pctl‖		
Andersen et al. (2018)	9 5–19†	1.10 (0.13–3.41) 2.5–97.5 pctl#	14.77 (11.87–18.93) 2.5–97.5 pctl#	TSH > 97.5 pctl fT4 < 2.5 pctl	TSH > 97.5 pctl fT4 normal	<11.87 = <2.5 pctl
Derakhshan et al. (2018a)	10 95% range: 6–14	0.11–3.48 2.5–97.5 pctl	11.6–19.4 2.5–97.5 pctl	Excluded from scope of Derakhshan et al. (2018a) study		
Nelson et al. (2018)	10 8–12*	0.07–2.55 2.5–97.5 pctl	12.53–22.78 2.5–97.5 pctl	TSH > 97.5 pctl fT4 < 2.5 pctl	TSH > 97.5 pctl fT4 normal	<12.53 = <2.5 pctl
Spann et al. (2020)	"early to mid-gestation"†	>0.13–3.63 >5–95 pctl	>0.82–1.32 ng/dL >5–95 pctl	TSH > 3.63 (>95 pctl)	TSH > 3.63 (>95 pctl) fT4: >0.82–1.32 ng/	≤0.86 ng/dL = ≤10 pctl

(*continued*)

Table 1. Continued.

Publication	GW and GW range	Median, reference range		Hypothyroidism		Hypothyroxinaemia
		TSH [mU/L]	fT4 [pmol/L; or noted]	Overt	Subclinical	fT4: pmol/L
				fT4 $\leq$ 0.82 ng/dL ($\leq$5 pctl)	dL (>5–95 pctl)	
Lain et al. (2020)	10.2*		tT4: 10.1 µg/dL (SD: 2.0) tT3: 21.2 µg/dL (SD: 3.8) fT4 index: 2.1 SD: 0.4)	Dichotomised as >2.5 or $\geq$ 2.5 / $\geq$ 4.0	Not defined	Not defined
Randomised clinical trials						
Lazarus et al. (2012)	12 and 3 days; range not specified†	0.15–3.65 2.5–97.5 pctl	8.4–14.6 2.5–97.5 pctl		TSH > 97.5 pctl and/or fT4 < 2.5 pctl	
Hales et al. (2018)		1.16	14.12			

*TPO-antibody positive mothers excluded or TPO-antibody positivity considered as potential confounder.
†No mention of exclusion (or non-exclusion) of TPO-antibody positive mothers.
‡This study also considered tT4; setting the normal range at 58–128 nmol/L further indicating that this relates to the normal population reference range × 1.5 "based on the recommendations of the Endocrine Society Clinical Practice Guideline."
¶This study also considered tT4, at GW 10, a median of 132.6 nmol/L (SD: 31.8) was recorded.
§Upper reference limit: 2.5 multiples of median (Männistö et al. 2011).
||Korevaar et al. (2016a) aimed at investigating "the shape of the association between TSH or fT4 concentrations and child IQ or brain MRI." During further evaluations, mothers with overt hypothyroidism and overt hyperthyroidism (TSH < 97.5 percentile and fT4 > 2.5 percentile) were excluded from the evaluation.
#GW- and method-specific reference ranges established for this cohort (Laurberg et al. 2016).
GW: gestational week; Pctl: percentile; fT4: free thyroxine; SD: standard deviation; TPO: thyroid peroxidase; TSH: thyroid stimulating hormone; tT3: total T3; tT4: total thyroxine.

increased TSH) and impaired child neurodevelopment (Table 2). Overall, a variety of different statistically significant outcomes were reported. These differences can be explained both by the varied definition of the respective cohorts (including age of the child) and by the diversity of neurodevelopmental parameters studied.

The studies by Korevaar et al. (2016a) and Jansen et al. (2019) deserve further consideration even though not yet conclusive in their findings (see Box 1 for details). In addition to child IQ, these studies included brain MRI as measurement of child neurodevelopmental outcomes. In brief, Korevaar et al. (2016a) recorded a statistically significant inverted U-shaped association of maternal fT4, but not TSH, with child total gray matter (at the age of 8 years). The U-shaped response curve suggests that there is an optimum range of fT4 concentration below or above which the amount of gray matter is decreased. By contrast, Jansen et al. (2019) recorded a statistically significant inverted U-shaped association of maternal TSH with child total gray matter (at the age of 9.9 years). (By analogy, this U-shaped response curve suggests there is an optimum TSH range below or above which the total gray matter is decreased.) As compared to the findings reported by Korevaar et al. (2016a), the association between maternal fT4 and child total gray matter did not remain statistically significant when adjusted for total brain volume in the study by Jansen et al. (2019). To investigate whether the discrepancy between the findings of the two studies could be caused by selection bias by MRI data availability, Jansen et al. (2019) performed several sensitivity analyses, which did suggest some potential selection effects that might have influenced the results.

Human evidence for impact of liver enzyme induction on serum T4 (and TSH) levels

For Objective 2, human studies were evaluated to collate information on the following sequence of three key events:

1. The impact of substance-mediated liver enzyme induction;
2. On human (maternal) serum thyroid hormone levels;
3. Leading to child neurodevelopment in case of *in utero* exposure.

Focus was on studies investigating the impact of antiepileptic drugs on (maternal) liver enzyme induction and thyroid function in humans and, ultimately, child neurodevelopment. Due to the paucity of information retrieved addressing the entire sequence of three key events, studies investigating the impact of non-pharmaceutical substances were also included in the evaluation if they addressed all three key events.

While the search queries were kept broad (key terms "liver," "hepatic") to ensure comprehensiveness of the database, the evaluation focused on the induction of hepatic UGTs as important enzymes for thyroid hormone

Table 2. Association between maternal serum fT4 and TSH levels and statistically significant neurodevelopmental outcomes in the offspring.

Findings in the pregnant mothers	Statistically significant findings in the offspring	Reference
Association between reduced maternal serum fT4 levels and statistically significant neurodevelopmental outcomes in the offspring		
fT4 < 10th percentile at GW 12	Impaired child psychomotor development at 10 months	Pop et al. (1999)
fT4 < 10th percentile at GW 13	Expressive language delay when evaluating 18- and 30-month-old children together	Henrichs et al. (2010)
fT4 < 5th percentile at GW 13	Decreased child mental scores at 14 months	Julvez et al. (2013)
fT4 < 11.4 pmol/L at GW 10.7	Repetition of school class in boys	Päkkilä et al. (2015)
fT4 < 3rd to <11th (and >88th to >97th percentiles) at GW 13.2	Lower child IQ at 6 years	Korevaar et al. (2016a)
Continuous range of fT4 levels	Inverted U-shaped association with child total gray matter and cortex volume (MRI) in subset of 646 children (no statistical significance within gestational normal-range fT4)	Korevaar et al. (2016a)
fT4 < 2.5th percentile	Autism spectrum disorders in girls	Andersen et al. (2018)
fT4 < 10th percentile	Fivefold increased risk of bipolar disorder with psychotic features; fourfold increased risk for offspring bipolar disorder in females (6 cases), but not in males (2 cases)	Spann et al. (2020)
Association between increased maternal serum TSH levels and statistically significant neurodevelopmental outcomes in the offspring		
Higher TSH at GW 13.3	Attention problems and aggressive behaviour when both the mother- and father-reported scores were combined using both the data from 1.5- and 3-year-old children	Ghassabian et al. (2011)
TSH > 3.1 mU/L at GW 10.7	Self-evaluated difficulties in mathematics in girls	Päkkilä et al. (2015)
Continuous range of TSH levels	Inverted U-shaped association with offspring total gray matter volume and with cortical gray matter volume	Jansen et al. (2019)
No statistically significant associations between altered maternal fT4 or TSH and neurodevelopmental outcomes in the offspring		
No association between maternal TSH or TPO antibody levels at GW 10 and child cognitive scores at 5.5 years		Williams et al. (2013)
Maternal thyroglobulin antibody levels associated with significantly lower scores on Perceptual Performance and Motor scales		Williams et al. (2013)
No associations between maternal thyroid function and (1) child educational attainment at 4.5 and 15 years; (2) child emotional and behavioural problems at 3.5, 6.75, 9 and 11 years; or (3) child ADHD at 7.5 and 15 years		Nelson et al. (2018) Fetene et al. (2018, 2020)
Maternal tT4 and TSH were not associated with child cognitive outcomes at 6 months, 3 or 7.7 years and maternal total T3, fT4, TPO antibody levels were further not associated with child cognitive outcomes at 7.7 years. Lain et al. (2020) did though note that the cohort size of 514 mother–child pairs may have been too small to observe significant differences		Oken et al. (2009) Lain et al. (2020)

fT4: free thyroxine; GW: gestational week; IQ: intelligence quotient; MRI: magnetic resonance imaging; TSH: thyroid stimulating hormone.

Box 1. Associations between maternal thyroid function and child brain MRI reported by Korevaar et al. (2016a) and Jansen et al. (2019).

Korevaar et al. (2016a)

A total of 3,839 mother–child pairs of the Generation R cohort were considered. It was strived to establish associations between maternal thyroid function and child neurodevelopmental outcomes across the entire range of maternal TSH and fT4 measurements.

Child IQ was measured in all children. Brain MRI analysis was conducted in a subset of 646 children (age: 8 years).

During statistical analysis, the shape of association between TSH or fT4 and child IQ or brain MRI outcomes was assessed using ordinary least-squares linear regression models with restricted cubic splines with three knots at the 10th, 50th, and 90th percentiles, building multiple linear regression models accordingly. Additionally, standard multivariate linear regression models were created with quadratic terms to provide effect estimates.

Child total gray matter volume and cortex volume were positively associated with child IQ ($p = 0.0044$). There was a statistically significant inverted U-shaped association between maternal fT4 levels and child total gray matter and cortex volume ($p = 0.0062$ and 0.0011, respectively). This association remained similar after excluding women with overt hypothyroidism and overt hyperthyroidism (Table 1). Within gestational normal-range fT4 (10.2–22.4 pmol/L; $n = 613$), the overall association of maternal fT4 with mean child gray matter or cortex volume did not remain statistically significant.

There was a non-significant inverted U-shaped association between maternal fT4 and child IQ when all 3,839 mother–child pairs were evaluated. In a subset of 598 mother–child pairs with data for both IQ and brain MRI, maternal fT4 concentration was not associated with child IQ.

Maternal TSH levels were not associated with child IQ. Associations between maternal TSH and child total gray matter volume and cortex volume disappeared after correction for total brain volume.

All results were independent of maternal human chorionic gonadotropin, TPO antibodies, or child TSH and fT4 concentrations.

Jansen et al. (2019)

Similar setting using mother–child pairs from the Generation R cohort as applied by Korevaar et al. (2016a). Here, brain MRI was conducted in 1,981 children (median age: 9.9 years). Multivariable linear regression models were used to study the association of maternal TSH and fT4 with brain outcomes.

There was an inverted U-shaped association of maternal TSH and fT4 with offspring total gray matter volume ($p = 0.007$) and with cortical gray matter volume ($p = 0.022$). When the analyses were adjusted for total brain volume, only TSH but not fT4 remained statistically significant. The association between maternal TSH and child total gray matter volume was most evident for TSH levels within the first 14 weeks of pregnancy.

metabolism (and as stands in line with the focus of Appendix A of EFSA and ECHA (2018) Endocrine Disruptor Guidance). UGTs are phase II enzymes that catalyze thyroid hormone glucuronidation, thereby increasing their solubility for renal and biliary excretion (Strolin Benedetti et al. 2005). Other pathways, including sulfation, deiodination, ether bond cleavage and/or oxidative deamination, as well as transport processes, also contribute to thyroid hormone clearance in humans (DeGroot and Jameson 2001; Wong et al. 2005). Relevant aspects of such metabolic processes are further discussed in the planned second review on thyroid-related AOPs which include neurodevelopmental adverse outcomes.

Evidence for impact of maternal exposure to liver enzyme-inducing substances on child neurodevelopment

The literature search yielded one article (Builee and Hatherill 2004) addressing all three key events of Objective 2. Builee and Hatherill (2004) is not a research article, but a literature review. It did not consider antiepileptic drugs, but the impact of polyhalogenated aromatic hydrocarbons on UGT-induction in pregnant mothers leading to low maternal serum thyroid hormone levels and consequently child neurodevelopmental impairment. Importantly, Builee and Hatherill (2004) only discussed the mechanisms by which liver enzyme-induced T4 reductions leading to neurodevelopmental impairment might occur upon exposure to these substances, but provided no evidence that the sequence of key events actually occurs. Specifically, evidence was only provided for either substance-mediated liver enzyme induction leading to altered serum T4 levels or altered serum T4 levels leading to child neurodevelopment, but made no attempt to link the two processes. Further, the scientific literature cited in regard to liver enzyme induction exclusively relates to *in vitro* assays or *in vivo* rodent studies and takes no account of potential species differences in the control and regulation of hepatic and thyroid hormone control. Of note, the different polyhalogenated aromatic hydrocarbons considered by Builee and Hatherill are no longer produced or marketed in Europe or North America (Council 1985; EP and Council 2003; US EPA 2017).

Another literature review (Verrotti et al. 2014) addressed the possible link between maternal exposure to antiepileptic drugs and developmental neurotoxicity in the offspring, however without consideration of liver enzyme induction. This review had a focus on animal studies supplemented by information from *in vitro* developmental neurotoxicity assays, but also summarised human studies addressing if *in utero* exposure to antiepileptic drugs affects child neurodevelopment. For antiepileptic drugs that are known to induce liver enzymes, some human studies reported an adverse impact of *in utero* exposure on offspring IQ whereas others did not (Verrotti et al. 2014).

From amongst the human studies summarised by Verrotti et al. (2014), the study by Reinisch et al. (1995) was selected for further evaluation as it reported an adverse impact of *in utero* exposure to phenobarbital on cognitive function (parameter: IQ levels in adult men). Two independent double-blind studies were presented. The second study was conducted with the aim of replicating the initial findings using an independent study population and a different standard measure of intelligence. Generally, both studies presented by Reinisch et al. (1995) indicated that *in utero* exposure to phenobarbital resulted in reduced IQ in adult men. With respect to Objective 2 of the present review, Reinisch et al. did not investigate if and how phenobarbital exposure affected liver enzyme levels as an initiating event for increased thyroid hormone clearance. Further, the phenobarbital exposure groups were formed based upon the indication of phenobarbital treatment in the mothers' medical records, but there is no mention of why phenobarbital was prescribed to the mothers, and clinical information from the mothers was scant. Even though phenobarbital has a limited list of indications, the paucity of clinical information does not preclude that the mothers might have exhibited a specific pathophysiological feature that could also account for the cognitive findings of their children. Further, phenobarbital is, in itself, a sedative able to cross the placenta (De Carolis et al. 1992). Therefore, a clear distinction between a direct and indirect, thyroid hormone-mediated effect on the child nervous system cannot be made.

In summary, the reviews by Builee and Hatherill (2004) and Verrotti et al. (2014), just as the exemplary epidemiological study (Reinisch et al. 1995), did not include information of relevance for Objective 2. They did not indicate whether or not liver enzyme activities were actually induced in these studies, or whether there was any consequential change in the maternal serum T4, or, ultimately, in child neurodevelopment.

Evidence for impact of exposure to substances, which induce liver enzymes, on serum T4 (and TSH) levels

Due to the paucity of information retrieved, this part of the review considered a modification of the first two key events listed above with respect to Objective 2, i.e. (1) the impact of substance-mediated liver enzyme induction; (2) on serum thyroid hormone levels in human volunteers or patients (regardless of age or gender, but excluding exposure via lactation).

Hence, this part of the review did not consider the impact of *in utero* substance exposure on child neurodevelopment. Following these specifications, one meta-analysis and twelve human studies were retrieved and evaluated:

Zhang et al. (2016) is a meta-analysis of 35 studies encompassing 997 patients, which generally showed that antiepileptic drug treatment of epileptic patients induces a significant decrease in fT4 levels and an increase in TSH levels. The meta-analysis, just as the underlying studies, did not specifically consider liver enzyme induction. Instead, Zhang and coworkers very generally discussed evidence by others indicating that, by inducing hepatic microsomal enzyme systems, antiepileptic drugs likely accelerate the metabolism of thyroid hormones. With respect to human UGT, Zhang et al. (2016) denote that high levels of UGT were observed in some studies upon antiepileptic drug exposure. Eirís-Puñal et al. (1999) and Shorvon (2000) are cited to substantiate this statement. However, these two papers do not mention UGT induction (Box 2).

Twelve human studies were retrieved that addressed the impact of antiepileptic drug treatment on T4 levels (and some also on TSH levels) and included measurements of liver function (see Supplementary Information SI-3 for further details). None of these human studies included assessments of UGT. Instead, the following direct or indirect parameters of liver function were measured:

- Antipyrine clearance (measure of the induction of cytochrome P450/monooxygenases): Seven publications (Ohnhaus et al. 1981; Ohnhaus and Studer 1983; Connell et al. 1984; Perucca et al. 1984; Patsalos et al. 1990; Larkin et al. 1991; Isojärvi et al. 1995);
- Urinary excretion of D-glucaric acid or 6-β-hydroxy cortisol (indirect markers of cytochrome P450 induction): Four publications (Ohnhaus et al. 1981; Ohnhaus and Studer 1983; Perucca et al. 1984; Larkin et al. 1991);
- Gamma glutamyl transferase: Four publications (Luoma et al. 1980; Ohnhaus et al. 1981; Ohnhaus and Studer 1983; Isojärvi et al. 2001);
- Serum lipid profiles: Two publications (Attilakos et al. 2007; Elger et al. 2016);
- Serum bilirubin: Luoma et al. (1980);
- Thyroxine binding globulin: Ohnhaus et al. (1981);
- Thyroxine binding globulin and transthyretin: Larsen et al. (1970).

Reviews addressing the impact of exposure to liver enzyme-inducing substances on serum T4 (and TSH) levels

Eight reviews were retrieved that addressed how treatment with antiepileptic drugs leads to liver enzyme induction thereby affecting serum T4 (and TSH) levels (Curran and DeGroot 1991; Capen 1992; Anderson 2004; Strolin Benedetti et al. 2005; Ennulat et al. 2010; Ghassabian and Trasande 2018; Mughal et al. 2018). Relevant epidemiological information from these reviews is summarised in the Supplementary Information SI-4.

Overall, the eight reviews highlight:

1. Species differences between rats and humans regarding the impact of liver enzyme induction on T4 and TSH levels; and
2. The paucity of information from human studies related to the substance-mediated induction of UGTs.

Importantly, all of these reviews addressed substance-mediated liver enzyme induction in animal studies or *in vitro* studies. As such, they are only suggestive that similar consequences might occur in humans.

Discussion

Based upon the human evidence summarised above, this section aims at answering the questions raised in the Introduction with respect to Objective 1 (Evaluate human evidence for how maternal thyroid hormone levels and child neurodevelopmental effects may be linked) and Objective 2 (Evaluate human evidence for a link between (maternal) substance-mediated UGT induction and increased serum T4 clearance and, ultimately, child neurodevelopmental impairment). In addition, the scientific evidence supporting Appendix A of EFSA and ECHA (2018) Endocrine Disruptor Guidance and the European Commission (2017a) Thyroid Disruption Workshop Report is further discussed. The conclusions on Objectives 1 and 2 are then used to further pursue the overarching goal to identify parameters in human studies that appear most relevant for toxicological assessments.

Objective 1: human evidence for how maternal thyroid hormone levels and child neurodevelopment may be linked

Do internationally accepted normal values (ranges) for serum TSH or thyroid hormone levels in pregnant women exist? If not, are they needed e.g. for the interpretation of human studies?

Generally, internationally accepted generic ranges for "normal," "low," or "high" TSH, T4 or T3 levels do not exist (WHO 2012; Medici et al. 2015; Alexander et al. 2017; Jonklaas and Razvi 2019). Instead, population-based reference ranges in TPO-antibody negative women are propagated (see e.g. Alexander et al. 2017), with criteria for such ranges having been applied differently across epidemiology studies (see below). If population-based reference ranges are not available for TSH, the usage of fixed cutoffs has been recommended. For example, in 2011, the American Thyroid Association set 2.5 and 3.0 mU/L TSH during the first and second trimesters of pregnancy, respectively, as such fixed upper cutoffs (Stagnaro-Green et al. 2011).

Maraka et al. (2017) discussed the clinical implications of these upper limits:

> "According to these diagnostic criteria, subclinical hypothyroidism [...] is estimated to affect up to 15% of pregnancies in the US and 14% in Europe. This represents a fivefold increase in prevalence compared with the 2–3% prevalence of subclinical hypothyroidism

Box 2. References to the statement in Zhang et al. (2016) that high levels of UGT were observed in some studies upon antiepileptic drug exposure.

Eirís-Puñal et al. (1999) evaluated serum thyroid hormone and thyroxine binding globulin levels in epileptic children under long-term antiepileptic drug therapy. They did not report measuring any other liver parameters apart from thyroxine binding globulin. Children receiving carbamazepine or valproate (but not those receiving phenobarbital) had significantly reduced thyroxine binding globulin levels. Eirís-Puñal et al. discussed that thyroid hormone reductions might be caused by competitive displacement from thyroxine binding globulin or by increased hepatic clearance due to the enzyme-inducing activities of these drugs. This latter statement is not referenced.
Shorvon (2000) is a review on oxcarbazepine that mentions that this drug (other than carbamazepine) hardly induces hepatic cytochrome P450 and that it has little effects on thyroid or sex hormones.

before these criteria were established, raising the possibility of overdiagnosis of subclinical hypothyroidism and discussions […] about increasing the TSH cut-off limit to 4.0 mIU/L …"

Indeed, in 2017, the American Thyroid Association raised the upper cutoff for TSH in the first trimester of pregnancy to 4.0 mU/L (Alexander et al. 2017).

The European Commission (2017a) Thyroid Disruption Workshop Report also highlighted that there is no internationally accepted normal TSH range and discussed the implications thereof: "As there is no consensus on the upper normal range of TSH, there is a controversy about the definition and clinical relevance of subclinical hypothyroidism (Hamilton et al. 2008)" (p. 98 in European Commission 2017a).

The European Commission (2017a) Workshop Report indicates the following types of altered maternal serum TSH and/or T4 levels:

- **Overt hypothyroidism**: "TSH serum levels >10 mU/L and T4 levels below the reference range" (note: most likely, this definition relates to population-specific reference ranges);
- **Subclinical hypothyroidism** "is diagnosed as 'mild' with normal TSH serum levels between 4 and 10 mU/L and fT4 below the 10th percentile of the reference range. The criteria for 'severe' subclinical hypothyroidism vary somewhat. It is sometimes defined as TSH levels >10 mU/L. Other authors classify 'severe' hypothyroidism with TSH in the normal range and fT4 levels below the 5th percentile of the reference range" (p. 98 in European Commission 2017a).

Notably, this definition for "subclinical mild hypothyroidism" matches the usual definition for "isolated hypothyroxinaemia." By contrast, the term "subclinical severe hypothyroidism" appears contradictory in itself since severe findings should be overt. Since the European Commission (2017a) Workshop Report is the main reference used in Appendix A of EFSA and ECHA (2018) to substantiate the view that low maternal serum thyroid hormone levels *per se* result in child neurodevelopmental impairment, the shortcomings of the definitions for altered maternal serum TSH and/or T4 levels in the workshop report also compromise the statements made in Appendix A.

Normal fT4 ranges are usually expressed as specific intervals of values (percentiles) with the absolute fT4 values being study-specific. The *International Federation of Clinical Chemistry and Laboratory Medicine (IFCC)* has defined the reference interval (i.e. reference range) as the prediction interval which includes the central 95% of reference values or test results from well-defined healthy individuals (Ichihara et al. 2017). Thereby, values <2.5th percentile and >97.5th percentile lie outside the normal reference range and thus are "abnormal." This implies that, following the IFCC definition for reference interval, subclinical mild and severe hypothyroidism as defined by European Commission (2017a), and also hypothyroxinaemia as defined in a number of cohort studies included in the present narrative review (see below), include both "normal" and "abnormal" fT4 levels.

As mentioned above, "population-based" reference ranges for serum fT4 (and TSH) are propagated. "Population" can relate to e.g. a country, a region, or other pre-defined group of people. Thereby, population-based reference ranges are dependent on ethnicity, age, TPO-antibody status, and iodine status of the given population. Further, population-based fT4 and TSH reference ranges are assay- and laboratory-dependent (Oken et al. 2009; Korevaar et al. 2013, 2017; Medici et al. 2015; Spencer et al. 2015; Lain et al. 2017; Maraka et al. 2017; Derakhshan et al. 2018a, 2018b; Jonklaas and Razvi 2019; Midgley et al. 2019; Razvi and Hostalek 2019; Su et al. 2019; Yang et al. 2019). It is also noteworthy that fT4 measurements are influenced by the concentration of thyroxine binding protein present in the sample and that this concentration changes considerably during pregnancy (Moleti et al. 2014). Similarly, maternal fT4 levels vary physiologically during pregnancy (Laurberg et al. 2016) just as all endocrine systems are inherently flexible and can adapt to altered physiological states (Glinoer 1997).

It is very important to consider all of these aspects when striving to establish a population-specific normal serum fT4 range for pregnant women (as described e.g. in Roche 2009).

Indeed, the absolute values of the normal ranges of fT4 (and TSH) differed between the different human studies (e.g. fT4: 11–25 pmol/L (Henrichs et al. 2010), 8.9–12.3 pmol/L (Julvez et al., 2013), 15.0 ± 1.7 pmol/L (Williams et al., 2013); TSH: 0.35–5.5 mU/L (Oken et al., 2009), 0.03–2.5 mU/L (Henrichs et al., 2010), 1.2–3.1 mU/L (Williams et al., 2013); Table 1). This can at least partly be explained by the differences in the criteria to include mother–child pairs in the cohorts, in the specific time points (GW) of sampling of maternal blood within the first trimester of pregnancy, and in the hormone measurement assays applied (as well as in the laboratories conducting these assays).

In summary, due to the various factors affecting normal variability of fT4, it is not surprising that a single physiologically relevant normal range for fT4 could not be identified (Alexander et al. 2017). Addressing such issues requires wide base efforts within the scientific community to joins forces to generalise all available evidence and to provide evidence-based recommendations, e.g. as has been initiated within the *Consortium on Thyroid and Pregnancy* (Korevaar et al. 2016b).

As long as normal ranges for maternal fT4 (or TSH) are unavailable, transparency on the specific criteria applied to define cohorts and to evaluate study outcomes are key to ensure the relevance of findings and to enable at least limited between-study comparability. For example, as a starting point to the *Swedish Environmental Longitudinal, Mother and Child, Asthma and Allergy* study, Derakhshan et al. (2018a) pre-defined the reference ranges and determinants of thyroid function during early pregnancy in that cohort and highlighted that: "Population-based thyroid function references ranges are required to optimally diagnose thyroid diseases in a clinical setting and to perform studies on adverse outcomes" (Derakhshan et al. 2018a).

As regards further serum parameters of maternal thyroid function, maternal tT4 was only included in a few studies (Oken et al. 2009; Ghassabian et al. 2011; Lain et al. 2020) and the meta-analysis by Thompson et al. (2018). Total T3

was only considered in one study (Lain et al. 2020), whereas fT3 was not evaluated in any of the human studies (Table 1). The work by Oken et al. (2009) and Lain et al. (2020) focuses on parameters of neonatal thyroid function, whereas maternal parameters are only briefly addressed. Overall, the available data on maternal tT4, total T3 and free T3 do not allow for establishing any associations with parameters for child neurodevelopment. Therefore, these parameters are not considered further in the sections below.

Is it possible to identify thresholds for altered serum TSH or fT4 levels in pregnant women indicating an increased risk for child neurodevelopmental impairment?

Generally, no. While the human evidence evaluated in the present narrative review is overall in support of a link between maternal thyroid function and child neurodevelopmental outcomes, it is insufficient to establish quantitative boundaries for high TSH and/or low fT4 levels in the maternal serum beyond which an increased risk for child neurodevelopmental impairment may be expected.

This can, at least partly, be explained by pronounced differences in the study designs with respect to the evaluation of maternal thyroid function and, most substantially, to the wide spectrum of neurodevelopmental outcomes investigated in the children (Table 2). A crucial issue impairing the comparability of the reviewed studies is that confounders were differently defined and considered when selecting mothers for the respective cohorts (see above for population, assay- and laboratory-dependency of TSH and fT4 reference ranges). For example, some cohort studies excluded mothers with alcohol consumption before or during pregnancy, whereas others did not. Similarly, some cohort studies excluded TPO-antibody positive mothers, whereas others did not (Supplementary Information SI-2). TPO-antibody-positivity and altered levels of human chorionic gonadotropin appear to play prominent roles in the interpretation of thyroid dysfunction (Korevaar et al. 2017; Derakhshan et al. 2018b). Other confounders influencing foetal neurodevelopment, such as different genetic syndromes, maternal metabolic diseases (e.g. diabetes), drug use, smoking, pre-eclampsia, preterm birth, low birthweight, should also be considered when assessing the impact of maternal thyroid function on child neurodevelopmental outcomes (US EPA 2015; Ornoy et al. 2016).

In most studies, increased risks for neurodevelopmental impairment were reported in children born from women who exhibited isolated low serum fT4 levels during pregnancy. Only in three studies, such increased risks were reported for children born from mothers with increased TSH (i.e. Ghassabian et al. 2011; Päkkilä et al. 2015; Jansen et al. 2019). This is unsurprising since mothers with previous diagnoses of thyroid dysfunction (e.g. overt hypothyroidism) were generally excluded from the studies. Therefore, the majority of mothers with altered serum hormone parameters included in the cohorts were hypothyroxinaemic, i.e. they exhibited isolated low fT4.

When interpreting the clinical implications of isolated low maternal fT4, it is important to note that, following some of the cohort-specific thresholds, the range for "abnormal fT4" also included "normal" hormone levels as per IFCC definition for reference interval, i.e. 2.5th–97.5th percentiles (Ichihara et al. 2017). Pop et al. (1999) and Henrichs et al. (2010) defined mild and severe hypothyroxinaemia as fT4 levels undercutting the 5th and 10th percentiles, respectively. Julvez et al. (2013) defined hypothyroxinaemia as fT4 $<$ 5th percentile, and Ghassabian et al. (2011) as $<$10th percentile (Table 1). Hence, as per IFCC definition for reference interval, the definitions for hypothyroxinaemia applied in the different cohort studies include both abnormally reduced maternal fT4 levels (i.e. fT4 $<$ 2.5th percentile) as well as low, but normal maternal fT4 levels (i.e. fT4 between the 2.5th and 5th or 10th percentiles). Indeed, increased incidences of neurodevelopmental effects were also recorded for children of mothers exhibiting low fT4 levels, which were normal as per IFCC definition., Hence, the IFCC definition for "normality," which is based upon a statistical analysis of variance, does not necessarily coincide with the fT4 range indicating functional sufficiency. However, child neurodevelopment can also be impaired by mechanisms unrelated to the thyroid hormone system. It is most likely also for such reasons, that from the human evidence evaluated in the present review, it is not clear whether clear-cut quantitative boundaries for maternal serum fT4 (or TSH) levels indicating increased risk for child neurodevelopmental impairment can be determined.

This observation is supported by the findings reported by Korevaar et al. (2016a) and Jansen et al. (2019), who assessed the association between maternal hormone levels and child neurodevelopment across the entire range of measurements. The inverted U-shaped associations between maternal serum fT4 or TSH levels and child total gray mater recorded by Korevaar et al. (2016a) and Jansen et al. (2019), respectively, did not indicate clear-cut quantitative boundaries for either maternal fT4 or TSH below or above which there were *no* effects on child total gray matter or cortex volume.

In summary, the present review of available human studies did not allow the identification of a single physiologically relevant normal range for fT4 (or TSH), or quantitative boundaries for maternal fT4 reduction (and/or TSH increase) indicating an increased risk for neurodevelopmental impairment in the children. Nonetheless, it is biologically plausible that a quantitative boundary for low fT4 levels indicating such increased risks exists (Moog et al. 2017). The unavailability of this quantitative information indicates a pertinent research need.

Finally, the attempt to identify quantitative boundaries for maternal fT4 and/or TSH levels indicating increased risks for child neurodevelopmental impairment considered if a specific window of susceptibility could be identified within the first trimester of pregnancy, i.e. the period during which foetal dependency on maternal T4 for physiological neurodevelopment is greatest. Generally, this assessment was restricted by the circumstance that the vast majority of studies only included maternal blood sampling at one single timepoint. Only Pop et al. (1999) mentioned maternal blood sampling at two different timepoints (GW 12 and 32, i.e. during the first and the third trimester of pregnancy), and no publication included different timepoints of blood sampling during the

first trimester. However, Jansen et al. (2019), stratified the study results (mean child total gray matter and cortex volume at 10 years of age) by gestational age at maternal blood sampling (intervals: GW 8, 10, 12, 14, 16 and 18; overall median: GW 13.3). This stratification showed that the inverted U-shaped association of maternal TSH with child total gray matter volume and cortical gray matter volume was most evident at GW 8. By contrast, after about GW 14, maternal TSH was no longer associated with child brain morphology (Jansen et al. 2019). Similarly, a recent meta-analysis of the correlation between maternal iodine status and child IQ found a positive curvilinear association of maternal urinary iodine/creatinine ratio with mean verbal IQ (children assessed at 1.5–8 years) that was only present up to GW 14 (Levie et al. 2019a). In line with these findings, several of the observational studies deserve to be reassessed excluding mother–child pairs in which the maternal blood was sampled after the first trimester since they might yield different study outcomes.

Is there evidence for a correlation between altered serum TSH or thyroid hormone levels in pregnant women without a history of thyroid disease and increased risk for child neurodevelopmental impairment?

While the human evidence evaluated in the present review implies that such an association exists, it is currently unknown if this association is causal.

Generally, the clinical diagnosis of thyroid disease or thyroid dysfunction is difficult to make since thyroid-specific complaints overlap with numerous nonspecific symptoms such as fatigue and body weight gain (Midgley et al. 2019). Further, due to the differing normal ranges of fT4 (and all other hormones related to the thyroid hormone system), there is no clear-cut distinction between subclinical and overt thyroid disease or between "normality" and "abnormality" (see above). While all studies excluded mothers with diagnosed thyroid dysfunction, it is likely that serum TSH and fT4 levels of these mothers overlap with subclinical hypothyroidism or hypothyroxinemia as specifically addressed in the cohorts. Therefore, when reviewing the human studies, it is not possible to differentiate if only pregnant women with overt hypothyroidism (i.e. low fT4 and high TSH levels) have a higher risk for impaired neurodevelopment of their children, or if this risk is also present in case of isolated reduced fT4 (i.e. hypothyroxinaemia) or isolated TSH (i.e. subclinical hypothyroidism). Andersen (2019) found 12.5% of randomly assessed pregnant mothers to exhibit abnormal thyroid function in the early pregnancy, with subclinical thyroid function abnormalities expectedly being more frequent than overt disease.

Further, other than in animal toxicity studies, where the determination of adversity is often based upon histopathological evaluations, the observational human studies generally do not include histopathology of the thyroid gland. Therefore, they do not provide information on whether (and if so, on how) reduced serum fT4 levels might be correlated with histological alterations of the thyroid gland. If histological alterations of the thyroid gland were observed, this would strongly indicate that any concordant low serum fT4 levels were indeed sufficient to activate the hypothalamic–pituitary–thyroid axis. By contrast, low serum fT4 recorded in the absence of such histopathological alterations might also indicate an incidental finding or a transient and physiological adaptive response of the highly versatile thyroid hormone system.

Appendix A of EFSA and ECHA (2018) Endocrine Disruptor Guidance states: "It is known from the broad knowledge of biology (e.g. human clinical experience and epidemiological data) that a drop in T4 results in impaired pre- and postnatal neurological development." This statement is linked to Alshehri et al. (2015) as the sole reference. Alshehri et al. (2015) is a literature review on the role of transthyretin, a thyroid hormone serum binding protein, in neurobiology. Therefore, it is unclear why the statement on T4 and child neurodevelopment in Appendix A is linked to Alshehri et al. (2015). Interestingly, the review by Alshehri et al. (2015) also plays a pivotal role in the European Commission (2017a) Thyroid Disruption Workshop Report that in turn is used as the sole reference in Appendix A to substantiate the view that rats and humans are equally sensitive to thyroid disruption. In this workshop report, different statements on such equal sensitivity are linked to Alshehri et al. (2015) as the sole reference. The Supplementary Information SI-5 provides a detailed presentation and discussion of the sequence of references tracing back the evidence presented by Alshehri et al. (2015). This sequence of references clearly shows that the Alshehri et al. (2015) review does not provide any objective information on a link between maternal T4 reduction with, or without, history of thyroid disease and child neurodevelopmental outcomes, or on human versus rodent sensitivity to thyroid disruption.

Is there evidence that T4 supplementation in healthy pregnant women with high TSH and/or low T4 serum levels decreases the risk for neurodevelopmental impairment in their children?

The evaluated human evidence does not allow establishing if T4 supplementation in healthy pregnant women with isolated low serum T4 levels will decrease the risk for neurodevelopmental impairment in their children.

In the randomised clinical trial by Lazarus et al. (2012) and Hales et al. (2018), T4 supplementation (i.e. levothyroxine therapy) in healthy pregnant women with low serum fT4 values did not decrease the risk for reduced cognitive function in their children. However, most likely, treatment of suboptimal gestational thyroid function was initiated too late in pregnancy. While the trial excluded pregnant mothers if the pregnancy had exceeded GW 15, the effect of suboptimal fT4 concentrations is presumably most detrimental in the first trimester of pregnancy so that treatment should have been initiated even earlier.

Similar to the findings by Lazarus et al. (2012) and Hales et al. (2018), Casey et al. (2017) reported that treatment for subclinical hypothyroidism and hypothyroxinaemia beginning before GW 17 and at GW 18, respectively, did not result in significantly better cognitive outcomes in children up to

5 years of age (as compared to no treatment for those conditions). Hence, also in this trial, treatment may have been initiated too late in pregnancy to positively affect child neurodevelopmental outcomes (see also our discussion above related to "Is it possible to identify thresholds for altered serum TSH or fT4 levels in pregnant women indicating an increased risk for child neurodevelopmental impairment?", which shows that any associations between maternal thyroid hormone levels and child neurodevelopment were most pronounced at GW 8, and no longer observable after approximately GW 14).

In this respect, the recommendations of the European Thyroid Association (Lazarus et al. 2014) regarding subclinical hypothyroidism in pregnancy are noteworthy: " … there is no demonstrable effect of maternal levothyroxine treatment on child neurodevelopment in relation to maternal subclinical hypothyroidism or maternal hypothyroxinaemia." Nevertheless, for the potential benefit of the child, it is recommended: " … levothyroxine therapy may be considered in isolated hypothyroxinaemia detected in the first trimester because of its association with neuropsychological impairment in children" (Lazarus et al. 2014).

Can the most sensitive parameter(s) addressing child neurodevelopmental outcomes be identified?

No, not from the evidence collected in the present review.

Just as the definitions for maternal thyroid dysfunction differed between studies, the human studies applied a broad variety of behavioural, neurodevelopmental neurological and anatomical parameters and statistical analysis methodologies to evaluate child neurodevelopmental outcomes (Table 2 and Supplementary Information SI-2). Clearly, the methods to establish neurodevelopmental impairment are not standardised. The between-study differences in the applied parameters make it difficult, if not impossible, to identify the most sensitive (or most relevant) parameter indicating child neurodevelopmental impairment. Similarly, a WHO (2012) Report on possible early effects of endocrine disruptors on child health concluded: "Effects on cognitive function resulting from exposure to thyroid disrupting chemicals are extremely difficult to estimate. It is not yet clear which specific cognitive functions, or methods of testing, may be the most representative of thyroid function during development."

The evidence for neurodevelopmental impairment recorded in the human studies was mostly weak. Often, statistically significant effects were only recorded in a few of many parameters that were further assessed without consideration of the clinical relevance of findings, or of an *a priori* hypothesis (Ghassabian et al. 2011; Korevaar et al. 2016a; Andersen et al. 2018).

The findings by Korevaar et al. (2016a) and Jansen et al. (2019) suggest that brain MRI measurements in the child may yield relevant and precise measurable outcomes to relate with disruption of maternal thyroid homeostasis. However, the findings from these studies are not yet conclusive on a casual relation between maternal thyroid function and brain morphological alterations in the child.

In spite of the limitations of the available evidence, the findings from the human studies considered in the present narrative review generally stand in line with current scientific understanding of endocrinology and thyroid pathology indicating that maternal TSH and T4 levels play a crucial role in foetal neurodevelopment (DeGroot and Jameson 2001). Pregnant women diagnosed with thyroid disorders are treated clinically and monitored closely to prevent child neurodevelopmental impairment (Lazarus et al. 2014).

Objective 2: human evidence for impact of liver enzyme induction on maternal serum T4 (and TSH) levels and, ultimately, child neurodevelopmental impairment

Is it physiologically plausible that liver enzyme induction has an impact on maternal serum thyroid hormone levels and, ultimately, child neurodevelopment? If so, to which extent would such an effect be compensated for by the TSH feedback system?

Generally, none of the papers reviewed provided relevant information to establish a link between substance-mediated induction of UGT, as important phase II enzyme for thyroid hormone metabolism, leading to increased serum thyroid hormone clearance, let alone to establish a further link between low maternal serum thyroid hormone and child neurodevelopmental impairment.

Indeed, none of the human studies addressing the impact of exposure to liver enzyme-inducing substances on serum thyroid hormone levels included the measurement of UGT. UGT activity is generally measured in tissue or cell samples or sub-cellular fractions, such as hepatic microsomes (Chen et al. 2018), and liver biopsies are rarely if ever undertaken in human observational studies, i.e. without medical indication. While *in vitro* studies have enhanced an understanding of the isoenzymes involved in thyroid hormone metabolism (e.g. Visser et al. 1993 and Findlay et al. 2000), they do not inform on the complex interplay between liver enzyme induction and serum hormone levels in an intact organism.

Most likely, it is currently not possible to establish the specific sequence of key events in humans that are presumed to lead from substance-mediated liver enzyme induction to reductions in maternal serum thyroid hormone levels and ultimately to child neurodevelopmental impairment, let alone to establish quantitative aspects of the key event relationships or potential species differences thereof.

This poses difficulties in toxicological assessment, which becomes more complex when considering the lack of knowledge on the similarity between humans and rodents (the predominant species used in toxicological studies). Surprisingly, Appendix A of EFSA and ECHA (2018) Endocrine Disruptor Guidance states that: "In the absence of substance-specific data which provide proof of the contrary, humans and rodents are considered to be equally sensitive to thyroid-disruption (including cases where liver enzyme induction is responsible for increased thyroid hormone clearance)."

The European Commission (2017a) Thyroid Disruption Workshop Report is used as sole reference to substantiate this view. In this workshop report, the Fipronil case study is

"used as a starting point for exploring and discussing thyroid disrupting compounds that increase thyroid hormone clearance, via induction of liver enzymes, and the human relevance of this molecular initiating event" (p. 21 in European Commission 2017a).

The Fipronil case study in European Commission (2017a) includes one single human study, by Herin et al. (2011); see also Supplementary Information SI-6. Herin et al. (2011) assessed the thyroid function status of 159 workers involved in the production of fipronil. Fipronil and fipronil sulfone were detected in the serum samples of nearly all and all workers, respectively. Higher serum fipronil sulfone levels were significantly correlated with lower TSH concentrations (p value = 0.03). However, only 1 of the 159 workers exhibited serum TSH below the normal range of 0.4-4.4 mU/L. Herin et al. (2011) concluded: "This study did not show that chronic fipronil exposure was associated with an increase in thyroid test abnormalities." Herin et al. (2011) did not report assessing any specific parameters of liver function.

In conclusion, the study by Herin et al. (2011) does not serve to inform on human sensitivity to substance-induced liver enzyme induction leading to increased thyroid hormone clearance, and the corresponding statement in Appendix A of EFSA and ECHA (2018) is not supported by the evidence presented in the referenced European Commission (2017a) Workshop Report.

Summary of the findings and conclusions

Guidelines issued by authorities and agencies are essential to establish transparency on regulatory procedures to perform hazard and risk assessment within the respective applicable legal framework, and to ensure that such procedures are both repeatable and relevant, i.e. founded on the state-of-the-art. Therefore, such supporting documents should be evidence-based and clear in their assumptions and conclusions. This is not the case with respect to the current Appendix A of EFSA and ECHA (2018), or the supporting European Commission (2017a) Thyroid Disruption Workshop Report, as has been discussed above.

To contribute to filling the knowledge gaps identified in Appendix A, the present narrative review served to collate human evidence on maternal serum thyroid hormone reduction potentially leading to child neurodevelopmental impairment (Objective 1) and on substance-mediated liver enzyme induction potentially leading to increased thyroid hormone clearance and hence reduced serum thyroid hormone levels (Objective 2). The motivation for this work was the goal to identify relevant parameters in pregnant women and their offspring that should be reflected in toxicological assessments addressing whether substances have the potential to elicit thyroid hormone imbalances in pregnant women, and if so, whether this would lead to neurodevelopmental impairment in the child, including those scenarios in which liver enzyme induction is the initiating event.

Therefore, the human evidence is reassessed below in view of the overarching question raised in the title of this review: *Which parameters from human studies are most relevant for toxicological assessments?* Specifically, it is aimed to identify (1) serum parameters and (2) non-serum parameters that appear most relevant to identify maternal thyroid hormone imbalances; (3) behavioural, neurodevelopmental, neurological and/or anatomical parameters that appear most relevant to identify child neurodevelopmental impairment; and (4) liver function parameters that allow for identifying substance-mediated liver enzyme induction leading to increased serum thyroid hormone clearance. This appraisal also serves to determine research needs to establish a scientific foundation for the development of a toxicological testing strategy to assess a substance's potential to elicit such effects (see also Table 3).

Which serum parameters appear most relevant to identify maternal thyroid dysfunction?

Summary of the findings

Based upon the human evidence collected, serum TSH and fT4 are, by frequency, the predominant parameters measured for the assessment of maternal thyroid function. By comparison, tT4 and free or total T3 are measured much less frequently in human studies. However, the study designs of the different human studies varied greatly. Further, serum TSH and fT4 concentrations are population-, assay- and laboratory-dependent, and internationally accepted generic ranges for normal serum TSH or fT4 in pregnant women generally do not exist. For these reasons, it was not possible to identify quantitative boundaries for maternal serum TSH or fT4 levels indicating increased risks for child neurodevelopmental impairment. Similarly, it is currently unclear whether such boundaries are generalisable between assays, groups of individuals or even across species. It is even questionable if clear-cut thresholds exist at all – also because child neurodevelopment can be impaired by mechanisms unrelated to the thyroid hormone system. Nevertheless, the evidence is overall in support of a link between maternal thyroid function and child neurodevelopmental outcomes.

Conclusion

It is currently not possible to identify whether TSH and fT4 are the most relevant parameters reflecting maternal thyroid dysfunction potentially resulting in child neurodevelopmental impairment, or if other thyroid hormone-related parameters (or parameters unrelated to the thyroid hormone system) might be more relevant. This is particularly important since in rodent studies tT4 rather than fT4 or free or total T3 is the parameter generally measured. Evidence is also needed to distinguish between physiological adaptation of the thyroid hormone system and thyroid dysfunction.

Further research work appears merited to enhance the understanding of which maternal serum parameters are most relevant to detect thyroid dysfunction during pregnancy, and hence should be considered in toxicological assessments. Such research work should also consider the uncertainty related to the overlap between physiological adaptation of the thyroid hormone system and thyroid dysfunction.

Table 3. Low maternal serum thyroid hormone levels and child neurodevelopment: which parameters should be reflected in toxicological assessments? – summary of the findings from the present narrative review and recommendations for future research work.

Type of parameter	Summary of the findings	Identification of research needs
Maternal serum parameters to identify maternal thyroid dysfunction?	Maternal serum fT4 and TSH are by far the most frequently measured parameters; unclear if they are also the most relevant Quantitative boundaries for these parameters indicating increased risks for impaired child neurodevelopment are not identifiable; possibly, clear-cut boundaries are inexistent	Enhance the understanding of which serum parameters are most relevant to detect thyroid dysfunction during pregnancy Caveat: uncertainty related to overlap between physiological adaptation of the thyroid system and thyroid dysfunction
Maternal (non-serum) parameters to identify maternal thyroid dysfunction?	Human studies do not provide information on histological alterations of the thyroid gland or on tissue levels of thyroid hormones, e.g. in the brain Establishment of presence of correlation between altered serum fT4 or TSH in pregnant women without history of thyroid disease and impaired child neurodevelopment is currently not possible	Evaluate data from patients from which both blood samples and biopsies of the thyroid gland were taken (regardless of medial indication); strive to use such histopathological information to link human findings with rodent studies Caveat: thyroid hormone system might exhibit increased sensitivity during pregnancy Strive to establish if serum thyroid hormone levels reflect tissue thyroid hormone levels (in humans? In rodents?)
Neurodevelopmental parameters to identify impaired child neurodevelopment?	Broad variety of neurodevelopmental parameters; identification of most sensitive parameter indicating maternal thyroid-dysfunction-mediated impaired child neurodevelopment not possible Possibly, brain MRI provides useful, objective, and reproducible information on child brain function	Establish the best possible use of brain MRI to assess child neurodevelopmental outcomes and identify the specific maternal serum parameters that the MRI findings should be linked to
Liver function parameters to identify substance-mediated UGT induction?	There is no evidence from human studies to establish link between substance-mediated UGT induction leading to increased thyroid hormone clearance, let alone further to child neurodevelopmental impairment UGT generally measured in tissues, cells or sub-cellular fractions; liver biopsies are rarely if ever undertaken in human studies	Strive to identify and establish noninvasive, indirect markers of UGT activity in serum or urine to facilitate future investigations on the impact of exposure to liver enzyme-inducing substances on both UGT levels and the thyroid hormone system (in humans and rodents) – and, ultimately, child neurodevelopment

Which further (non-serum) parameters might be relevant to identify maternal thyroid dysfunction?

Summary of the findings

The human evidence evaluated in the present review did not permit the establishment of any correlation between alterations in serum TSH or fT4 levels in pregnant women, without a history of thyroid disease, and neurodevelopmental impairment in their children. This can partly be explained by the difficulty in clinically diagnosing thyroid disease. While all cohort and case–cohort studies excluded mothers with diagnosed overt thyroid dysfunction, it is likely that serum TSH and fT4 levels of these mothers overlap with those with subclinical hypothyroidism or hypothyroxinaemia that were specifically addressed in the studies. Further, observational human studies do not provide information on histological alterations of the thyroid gland, or on tissue levels of thyroid hormones (e.g. in the brain). Such information would be useful to determine whether maternal serum thyroid hormone imbalances are truly "adverse" or e.g. indicative of transient physiological adaptations of the thyroid hormone system, and/or whether they are sufficiently pronounced to result in child neurodevelopmental impairment.

Conclusion

It is currently not possible to identify further (non-serum) parameters indicating maternal thyroid disruption. Possibly, an evaluation of human studies including patients (regardless of age, gender and state of pregnancy) from which both blood samples and biopsies of the thyroid gland were taken (on account of any specific medical indication) might enhance the understanding on which serum parameters are linked to which types of alteration of the thyroid gland. For example, thyroid function is addressed during treatment of chronic hepatitis C (see e.g. Alimenti et al. 2006) so that biopsy samples might be available from this patient group. Such information might be useful to establish a link that connects human findings (serum hormone levels and histopathological findings) with the corresponding data from rodent toxicity studies. Notably, such evaluations should consider that the thyroid hormone system might exhibit increased sensitivity during pregnancy.

Similarly, it remains to be established whether serum thyroid hormone levels reflect tissue thyroid hormone levels (which can be addressed in rodent studies). This is particularly relevant in the case of the brain, given the specific coordinated regulation of deiodinases (mostly in the brain, liver and kidney) to compensate for altered hormone levels in the blood circulation (Bianco and da Conceição 2018).

Finally, in the long term, innovative technologies might provide opportunities to establish noninvasive parameters indicative of thyroid hormone imbalances, such as the study of specific microRNAs (Dong et al. 2015).

Which parameters appear most relevant to identify child neurodevelopmental impairment?

Summary of the findings

The human studies applied a broad variety of parameters and statistical analysis methodologies to evaluate child neurodevelopmental outcomes. Assessments addressed psychomotor and mental development, cognitive function (IQ), expressive vocabulary and educational attainment and further included brain MRI and clinical diagnoses of e.g. autism or attention deficit hyperactivity disorder. Most parameters were only used in single human studies. This vastly impaired between-study comparability. Due to this diversity, it was not possible to identify which behavioural, neurodevelopmental, or neurological parameter might be most sensitive at indicating neurodevelopmental effects mediated by maternal thyroid hormone imbalances. A note on potential promising tools arises from the MRI evaluations from Korevaar et al. (2016a) and Jansen et al. (2019). While divergent in their findings, these studies open the discussion that precise reproducible measures of brain imaging may be useful for toxicological assessments (including evaluations of species differences), provided that any recorded morphological alterations are truly caused by maternal thyroid dysfunction.

Conclusion

While MRI is a potentially useful, objective and reproducible parameter to inform on morphological alterations of the child brain, further research work is merited to determine how MRI, as morphometric methodology, may contribute to the assessment of the impact of maternal thyroid hormone imbalances on child neurodevelopment. Such research work should serve to identify the specific maternal serum parameter(s) that the MRI findings should be linked to, and to determine how usage of MRI can be combined with functional assessments of the brain to enable comprehensive evaluations of child neurodevelopmental outcomes. If so, this approach could also be applied in toxicological assessments using rodents.

Which liver function parameters appear relevant to identify substance-mediated liver enzyme induction leading to increased serum thyroid hormone clearance?

Summary of the findings

None of the human studies or literature reviews evaluated provided relevant information to establish a link between substance-mediated UGT induction, as important phase II enzyme for thyroid hormone metabolism, leading to increased thyroid hormone clearance and hence reduced (maternal) serum thyroid hormone levels, let alone to establish a further link to child neurodevelopmental outcomes. It appears unlikely that this information is currently available at all. UGT activity is generally measured in tissue or cell samples or sub-cellular fractions, and liver biopsies are rarely if ever undertaken in human studies.

Notably, a number of human studies addressing the impact of treatment with antiepileptic drugs on serum thyroid hormone levels did include noninvasive, indirect measures of phase I cytochrome P450/monooxygenases (i.e. antipyrine clearance and urinary excretion of D-glucaric acid or 6-β-hydroxy cortisol).

Conclusion

Currently, noninvasive liver function parameters to identify substance-mediated UGT induction leading to increased serum thyroid hormone clearance are unavailable. Therefore, such human data are also unavailable, for which reason it is most likely currently not possible to assess species differences (humans versus rodents) in the key event relationship between substance-mediated UGT induction and increased serum thyroid hormone clearance. Similarly, it currently does not appear possible to identify if such effects could be compensated for by the TSH feedback system in humans.

To facilitate future investigations on the impact of exposure to liver enzyme-inducing substances on both UGT levels and the thyroid hormone system (in humans and rodents) and, ultimately, child neurodevelopment, research work is recommendable to strive to identify and establish noninvasive, indirect markers of UGT activity in serum or urine (by analogy to the available noninvasive, indirect markers of phase I enzyme induction).

Final conclusion and next steps

Appendix A of EFSA and ECHA (2018) Endocrine Disruptor Guidance proposes that the human relevance of thyroid effects observed in animals "could be further investigated," specifying that this investigation should also serve to elaborate if the effects are mediated by UGT induction. Further, Appendix A proposes a testing scheme for this investigation. However, this scheme does not describe how the different parameters should be measured or how the generated data should be evaluated within a weight-of-evidence approach to reach a conclusion on whether, or not, the substance under investigation fulfills the legal criteria for an endocrine disruptor as implemented in European Commission (2017b, 2018).

To contribute to addressing these uncertainties, the ECETOC Special T4 Task Force, as a first part of its work, has conducted the present extensive, but narrative, review of the available human evidence. The scope of this review has been aligned with the focus of Appendix A of EFSA and ECHA (2018): Hypothyroidism and hypothyroxinaemia have been considered as indicative of maternal thyroid function, and substance-mediated UGT induction as specific initiating event. Other potentially relevant processes, such as hyperthyroidism, hyperthyroxinaemia, iodine deficiency or a substance's ability to cross the blood–brain barrier were not considered in Appendix A and as such not reviewed here.

Specifically, the present review has aimed to establish the necessary human evidence to define those parameters that should be addressed in toxicological assessments when evaluating if a substance has the potential to elicit maternal thyroid hormone imbalances, and if so, if it also has the potential to affect child neurodevelopment.

The outcome of this review does not allow identifying any specific parameters as being most relevant to assess maternal thyroid function (or UGT induction) or neurodevelopmental impairment in the child. It also does not allow establishing quantitative relationships between the key events leading from substance-mediated liver enzyme induction over increased serum thyroid hormone clearance to reduced serum thyroid hormone levels and, ultimately, child neurodevelopmental outcomes. Notwithstanding, the information gathered appears representative, and it is considered unlikely that an expansion of the database, e.g. to include unpublished pharmaceutical studies, would alter the present findings. Instead, based upon the identified knowledge gaps, potential future research work has been suggested that might serve to overcome the limitations of the current human database (Table 3).

In the next part of its work, the ECETOC Special T4 Task Force will assess the suitability of the available *in vivo* rodent toxicity test methods and *in vitro* assays in predicting thyroid hormone imbalances in humans and in identifying critical key events of thyroid-related AOPs leading to neurodevelopmental adverse outcomes, including possible species differences in such key events and the corresponding key event relationships. In spite of the scientific limitations of the available human evidence, the outcome of the present review will provide important insight for this further work.

Ultimately, the work of the ECETOC Special T4 Task Force shall contribute to the development and establishment of a science-based toxicity testing strategy, to reliably predict a substance's potential to impair child neurodevelopment via maternal thyroid hormone imbalances, and, possibly, liver enzyme induction as initiating event.

Acknowledgements

We would like to thank the members of the ECETOC Scientific Committee for their critical comments. Also, we would like to thank all members of the ECETOC Special T4 Task Force as well as all participants of the November 2019 Extended Task Force Meeting for valuable discussions that contributed to the basis for this review. We are indebted to Olivier de Matos, Secretary General of ECETOC, and his team at ECETOC (Alice Brousse, Andreea Cuciureanu, Lisa Wingate, Virginie van der Steeg) for organisational and technical assistance to the ECETOC Special T4 Task Force and for support in preparing and holding the November 2019 Extended Task Force Meeting. A special thank you for the preparation and holding of that meeting also goes to Ursel Blum and Christine Gahn (BASF SE, Germany). The authors also extend a note of appreciation to the three reviewers selected by the Editor and anonymous to the authors of the paper. The comments of these reviewers were very useful in improving the manuscript.

Declaration of interest

This manuscript relates to work undertaken by the European Centre for Ecotoxicology and Toxicology of Chemicals (ECETOC; www.ecetoc.org) Special Thyroxine (T4) Task Force. ECETOC is a scientific organisation which provides a collaborative space for scientists from industry, academia and governments. Its mission is to develop and promote practical, trusted and sustainable solutions to scientific challenges which are valuable to industry, as well as to the regulatory community and society in general. ECETOC is financed by its membership, which are the leading companies with interests in the manufacture and use of chemicals, biomaterials and pharmaceuticals (http://www.ecetoc.org/ecetoc-membership/member-companies/). Within the ECETOC Task Forces, Task Force members work within their regular working hours, but do not receive compensation by ECETOC. This manuscript also refers to discussions from an Extended Task Force Meeting held in Ludwigshafen, Germany. The meeting venue and catering, as well as transfers in Ludwigshafen, were sponsored by BASF SE (see below for company details). ECETOC provided travel and accommodations support for the non-industry participants of the meeting, i.e. UGS (train fare), JAP (economy air fare), and GS (car mileage). All coauthors, with the exception of UGS, a freelance scientific writer, participated at the meeting and contributed to the preparation of the manuscript without compensation.

The coauthors of this manuscript consist of Task Force members (AA, PAB, AC, NH, SJ, SM, SM-K, VS), Stewards from the ECETOC Scientific Committee (BvR, GS), a neuroendocrinologist who attended the Extended Task Force Meeting (JAP), and the scientific writer (UGS). The views expressed in this article are solely those of the coauthors and may not represent those of the sponsoring organisations.

UGS was hired by ECETOC to assist in the preparation of this manuscript. This included payment of working hours and reimbursement of accommodation and travel expenses (see above) for participation at the Extended Task Force Meeting (purpose: note taking).

AA and NH are employed by Bayer AG. Bayer AG markets products (or previously marketed products) containing some of the chemicals included in this paper. Further, the Bayer portfolio includes substances that may have to be tested for their potential to cause maternal thyroid disruption and subsequent developmental neurotoxicity. This manuscript was submitted to in-house review in different Bayer AG company divisions, but no changes were requested. AA's responsibilities at Bayer AG include being the Head of Epidemiology Women's Healthcare. NH's responsibilities include being the Bayer Crop Science Division Environment Safety Ecotoxicology Terrestrial Vertebrates Team Leader. Up until January 2020, NH was Chair of the ECETOC T4 Task Force, and she chaired the Extended Task Force Meeting. AA's and NH's expenses for the attendance at that meeting were paid by Bayer AG; they received no funding in cash or kind for their contribution to this manuscript.

PAB and AC are employed by Syngenta, an international agribusiness that markets crop protection chemicals and seeds. Syngenta markets products (or previously marketed products) containing some of the chemicals included in this paper. Further, the Syngenta portfolio includes substances that may have to be tested for their potential to cause maternal thyroid disruption and subsequent developmental neurotoxicity. This manuscript was subjected to the usual internal peer-review process in Syngenta, in this case by the Global Head of Human Safety, but no changes were requested. PAB's responsibilities within Syngenta are to provide strategic scientific advice on product safety issues to the company's Product Safety, Business Sustainability and Crop Protection Development organisations. AC's responsibilities include providing scientific support to research and development activities and to regulatory toxicology projects. PAB's and AC's expenses for the attendance at the Extended Task Force Meeting were paid by Syngenta; they received no funding in cash or kind for their contribution to this manuscript.

SJ is employed by Albemarle Europe SRL. Albemarle markets products (or previously marketed products) containing some of the chemicals included in this paper. Further, the Albemarle portfolio includes substances that may have to be tested for their potential to cause maternal thyroid disruption and subsequent developmental neurotoxicity. SJ's responsibilities include heading the corporate toxicology department of Albemarle Corporation, worldwide management of regulatory toxicology and risk assessment of the chemicals produced by Albemarle Corporation. SJ's expenses for the attendance at the Extended Task Force Meeting were paid by Albemarle; she received no funding in cash or kind for her contribution to this manuscript.

SM is employed by The Dow Chemical Company. The issues of hazard identification and risk assessment of thyroid-active compounds, and how these are assessed by regulatory/other agencies, impact substances of interest to the corporation. An in-house review of this manuscript was conducted by additional scientists of The Dow Chemical Company, but no changes were requested. SM's role is focused on Dow's science

strategy and testing, which includes the assessment of endocrine-active compounds. SM received no funding in cash or kind for her contribution to this manuscript. SM's expenses for the attendance at the Extended Task Force Meeting were paid by Dow.

BvR, SMK and VS are employed by BASF SE, Ludwigshafen, Germany. BASF SE produces a very wide range of chemicals including some of those mentioned in this paper and/or substances that may have to be tested for their potential to cause maternal thyroid disruption and subsequent developmental neurotoxicity. This paper underwent the normal BASF SE review process of the Global Product Safety Department of Agricultural Solutions, but no changes were requested. BvR is Senior Vice President of the BASF SE Experimental Toxicology and Ecology Department, which is independent of any business unit and ISO 17020 certified. BvR is an Associate Professor of Reproduction Toxicity of the University of Wageningen, Netherlands, and the Chairman of the ECETOC Scientific Committee. SMK's responsibilities within BASF SE include being regulatory toxicologist for agrochemicals, and she is the current Chair of the ECETOC T4 Task Force. VS is Principal Scientist and Head of the Clinical Pathology Laboratory at BASF SE. BvR, SMK and VS received no funding in cash or kind for their contribution to this manuscript, any travel expenses were covered entirely by BASF.

JAP is full professor at the School of Medicine, University of Minho in Braga, Portugal, and Vice-President of the Medical School. Her main scientific interests are on the areas of neuroendocrinology, thyroid hormones, iron, choroid plexus/cerebrospinal fluid barriers in health and neuropsychiatric diseases, and genetics of neuropsychiatric diseases. Due to her renowned expertise in these areas, ECETOC invited JAP to attend the Extended Task Force Meeting. Thereafter, JAP contributed to the writing of this manuscript. JAP received no funding in cash or kind for participation at the meeting or contribution to this manuscript other than coverage of travel expenses and accommodation. She effectuated all work related to this manuscript in the course of her regular working hours.

GS is an occupational epidemiologist with main scientific interests on the potential long-term adverse health effects of human exposure to chemicals. GS has worked at the Maastricht University, Netherlands, from 1982 to 2004 and from 2014 to 2019, and at the Dow Chemical Company from 2004 to 2013. He now has an honorary appointment at the Maastricht University, and is member of the ECETOC Scientific Committee. GS has received no funding in cash or kind for his contributions to the manuscript or his participation at the Extended Task Force Meeting other than the travel expenses and accommodation.

Finally, this manuscript was reviewed by the ECETOC Scientific Committee consisting of representatives of academia and industry (http://www.ecetoc.org/about-ecetoc/scientific-committee/). This review yielded very few and only minor comments.

Supplemental material

Supplemental material for this article is available online here.

ORCID

Alex Asiimwe http://orcid.org/0000-0002-7633-0220
Nina Hallmark http://orcid.org/0000-0002-4210-628X
Sue Marty http://orcid.org/0000-0002-5404-8515
Joana A. Palha http://orcid.org/0000-0002-8866-368X
Volker Strauss http://orcid.org/0000-0001-7363-9663
Bennard van Ravenzwaay http://orcid.org/0000-0002-4264-9217

References

Alexander EK, Pearce EN, Brent GA, Brown RS, Chen H, Dosiou C, Grobman WA, Laurberg P, Lazarus JH, Mandel SJ, et al. 2017. 2017 Guidelines of the American Thyroid Association for the diagnosis and management of thyroid disease during pregnancy and the postpartum. Thyroid. 27:315–389.

Alimenti A, Forbes JC, Oberlander TF, Money DM, Grunau RE, Papsdorf MP, Maan E, Cole LJ, Burdge DR. 2006. A prospective controlled study of neurodevelopment in HIV-uninfected children exposed to combination antiretroviral drugs in pregnancy. Pediatrics. 118:e1139–e1145.

Alshehri B, D'Souza DG, Lee JY, Petratos S, Richardson SJ. 2015. The diversity of mechanisms influenced by transthyretin in neurobiology: development, disease and endocrine disruption. J Neuroendocrinol. 27:303–323.

Andersen SL, Andersen S, Vestergaard P, Olsen J. 2018. Maternal thyroid function in early pregnancy and child neurodevelopmental disorders: a Danish nationwide case-cohort Study. Thyroid. 28:537–546.

Andersen SL. 2019. Frequency and outcomes of maternal thyroid function abnormalities in early pregnancy. Scand J Clin Lab Invest. 79: 99–107.

Anderson GD. 2004. Pharmacogenetics and enzyme induction/inhibition properties of antiepileptic drugs. Neurology. 63:S3–S8.

Attilakos A, Garoufi A, Voudris K, Mastroyianni S, Fotinou A, Papadimitriou DT, Gavalakis N, Prassouli A, Katsarou E. 2007. Thyroid dysfunction associated with increased low-density lipoprotein cholesterol in epileptic children treated with carbamazepine monotherapy: a causal relationship? Eur J Paediatr Neurol. 11:358–361.

Bianco AC, da Conceição RR. 2018. The deiodinase trio and thyroid hormone signaling. Methods Mol Biol. 1801:67–83.

Boyages SC, Halpern JP. 1993. Endemic cretinism: toward a unifying hypothesis. Thyroid. 3:59–69.

Builee TL, Hatherill JR. 2004. The role of polyhalogenated aromatic hydrocarbons on thyroid hormone disruption and cognitive function: a review. Drug Chem Toxicol. 27:405–424.

Capen CC. 1992. Pathophysiology of chemical injury of the thyroid gland. Toxicol Lett. 64–65:381–388.

Casey BM, Thom EA, Peaceman AM, Varner MW, Sorokin Y, Hirtz DG, Reddy UM, Wapner RJ, Thorp JM Jr, Saade G, et al. 2017. Treatment of subclinical hypothyroidism or hypothyroxinemia in pregnancy. N Engl J Med. 376:815–825.

Chen A, Zhou X, Cheng Y, Tang S, Liu M, Wang X. 2018. Design and optimization of the cocktail assay for rapid assessment of the activity of UGT enzymes in human and rat liver microsomes. Toxicol Lett. 295: 379–389.

Chevrier J, Harley KG, Kogut K, Holland N, Johnson C, Eskenazi B. 2011. Maternal thyroid function during the second half of pregnancy and child neurodevelopment at 6, 12, 24, and 60 months of age. J Thyr Res. 2011:1–13.

Choksi NY, Jahnke GD, St Hilaire C, Shelby M. 2003. Role of thyroid hormones in human and laboratory animal reproductive health. Birth Defects Res B Dev Reprod Toxicol. 68:479–491.

Connell JM, Rapeport WG, Gordon S, Brodie MJ. 1984. Changes in circulating thyroid hormones during short-term hepatic enzyme induction with carbamazepine. Eur J Clin Pharmacol. 26:453–456.

Costeira MJ, Oliveira P, Santos NC, Ares S, Saenz-Rico B, de Escobar GM, Palha JA. 2011. Psychomotor development of children from an iodine-deficient region. J Pediatr. 159:447–453.

Council. 1985. Council Directive of 1 October 1985 amending for the sixth time (PCBs/PCTs) Directive 76/769/EEC on the approximation of the laws, regulations and administrative provisions of the Member States relating to restrictions on the marketing and use of certain dangerous substances and preparations (85/467/EEC). OJ L. 269:56–58.

Curran PG, DeGroot LJ. 1991. The effect of hepatic enzyme-inducing drugs on thyroid hormones and the thyroid gland. Endocr Rev. 12: 135–150.

De Carolis MP, Romagnoli C, Frezza S, D'Urzo E, Muzii U, Mezza A, Ferrazzani S, De Carolis S. 1992. Placental transfer of phenobarbital: what is new? Dev Pharmacol Ther. 19:19–26.

DeGroot LJ, Jameson JL. 2001. Endocrinology. Vol. I. Philadelphia (PA): WB Saunders Company.

Derakhshan A, Shu H, Broeren MAC, de Poortere RA, Wikström S, Peeters RP, Demeneix B, Bornehag CG, Korevaar TIM. 2018a. Reference ranges and determinants of thyroid function during early pregnancy: the SELMA study. J Clin Endocrinol Metab. 103:3548–3556.

Derakhshan A, Korevaar TIM, Taylor PN, Levie D, Guxens M, Jaddoe VWV, Nelson SM, Tiemeier H, Peeters RP. 2018b. The association of maternal

thyroid autoimmunity during pregnancy with child IQ. J Clin Endocrinol Metab. 103:3729–3736.
Dickhoff WW, Darling DS. 1983. Evolution of thyroid function and its control in lower vertebrates. Am Zool. 23:697–707.
Dong H, You SH, Williams A, Wade MG, Yauk CL, Thomas Zoeller R. 2015. Transient maternal hypothyroxinemia potentiates the transcriptional response to exogenous thyroid hormone in the fetal cerebral cortex before the onset of fetal thyroid function: a messenger and microRNA profiling study. Cereb Cortex. 25:1735–1745.
Drover SSM, Villanger GD, Aase H, Skogheim TS, Longnecker MP, Zoeller RT, Reichborn-Kjennerud T, Knudsen GP, Zeiner P, Engel SM. 2019. Maternal thyroid function during pregnancy or neonatal thyroid function and attention deficit hyperactivity disorder: a systematic review. Epidemiology. 30:130–144.
ECHA. 2018 Feb 15. How chemicals can affect the health of developing children. ECHA Newsletter. p. 24.
EFSA, ECHA. European Food Safety Authority and European Chemicals Agency with the technical support of the Joint Research Centre, Andersson N, Arena M, Auteri D, Barmaz S, Grignard E, Kienzler A, Lepper P, Lostia AM, Munn S, Parra Morte JM, et al. 2018. Guidance for the identification of endocrine disruptors in the context of Regulations (EU) No 528/2012 and (EC) No 1107/2009. ECHA-18-G-01-EN. EFSA J. 16:1661–1170.
Eirís-Puñal J, Del Río-Garma M, Del Río-Garma MC, Lojo-Rocamonde S, Novo-Rodríguez I, Castro-Gago M. 1999. Long-term treatment of children with epilepsy with valproate or carbamazepine may cause subclinical hypothyroidism. Epilepsia. 40:1761–1766.
Elger CE, Rademacher M, Brandt C, Elmoufti S, Dedeken P, Eckhardt K, Tennigkeit F, De Backer M. 2016. Changes in hormone and lipid levels in male patients with focal seizures when switched from carbamazepine to lacosamide as adjunctive treatment to levetiracetam: A small phase IIIb, prospective, multicenter, open-label trial. Epilepsy Behav. 62:1–5.
Ennulat D, Walker D, Clemo F, Magid-Slav M, Ledieu D, Graham M, Botts S, Boone L. 2010. Effects of hepatic drug-metabolizing enzyme induction on clinical pathology parameters in animals and man. Toxicol Pathol. 38:810–828.
EP and Council. 2003. Directive 2003/11/EC of the European Parliament and of the Council of 6 February 2003 amending for the 24th time Council Directive 76/769/EEC relating to restrictions on the marketing and use of certain dangerous substances and preparations (pentabromodiphenyl ether, octabromo- diphenyl ether). OJ L. 42:45–46.
European Commission. 2017a. Supporting the organisation of a workshop on thyroid disruption – final report. Framework Contract ENV.A.3/FRA/2014/0029 on implementation of the Community strategy on endocrine disruptors. Written by: Brunel University London and DTU National Food Institute Denmark.
European Commission. 2017b. Commission Delegated Regulation (EU) 2017/2100 of 4 September 2017 setting out scientific criteria for the determination of endocrine-disrupting properties pursuant to Regulation (EU) No 528/2012 of the European Parliament and Council. OJ L. 301:1–5.
European Commission. 2018. Commission Regulation (EU) 2018/605 of 19 April 2018 amending Annex II to Regulation (EC) No 1107/2009 by setting out criteria for the determination of endocrine disrupting properties. OJ L. 101:33–36.
Fan X, Wu L. 2016. The impact of thyroid abnormalities during pregnancy on subsequent neuropsychological development of the offspring: a meta-analysis. J Matern Fetal Neonatal Med. 29:3971–3976.
Fetene DM, Betts KS, Alati R. 2017. Mechanisms in endocrinology: maternal thyroid dysfunction during pregnancy and behavioural and psychiatric disorders of children: a systematic review. Eur J Endocrinol. 177: R261–R273.
Fetene DM, Betts KS, Alati R. 2018. Maternal prenatal thyroid function and offspring ADHD: findings from the ALSPAC cohort. J Nerv Ment Dis. 206:859–864.
Fetene DM, Betts KS, Scott JG, Alati R. 2020. Maternal prenatal thyroid function and trajectories of offspring emotional and behavioural problems: findings from the ALSPAC cohort. Eur Child Adolesc Psychiatry. 29:871–879.
Findlay KA, Kaptein E, Visser TJ, Burchell B. 2000. Characterization of the uridine diphosphate-glucuronosyltransferase-catalyzing thyroid hormone glucuronidation in man. J Clin Endocrinol Metab. 85:2879–2883.
Fisher DA, Klein AH. 1981. Thyroid development and disorders of thyroid function in the newborn. N Engl J Med. 304:702–712.
Ghassabian A, Bongers-Schokking JJ, Henrichs J, Jaddoe VWV, Visser TJ, Visser W, de Muinck Keizer-Schrama SMPF, Hooijkaas H, Steegers EAP, Hofman A, et al. 2011. Maternal thyroid function during pregnancy and behavioral problems in the offspring: the Generation R study. Pediatr Res. 69:454–459.
Ghassabian A, Trasande L. 2018. Disruption in thyroid signaling pathway: a mechanism for the effect of endocrine-disrupting chemicals on child neurodevelopment. Front Endocrinol. 9:549–548.
Glinoer D. 1997. The regulation of thyroid function in pregnancy: pathways of endocrine adaptation from physiology to pathology. Endocr Rev. 18:404–433.
Haddow JE, Palomaki GE, Allan WC, Williams JR, Knight GJ, Gagnon J, O'Heir CE, Mitchell ML, Hermos RJ, Waisbren SE, et al. 1999. Maternal thyroid deficiency during pregnancy and subsequent neuropsychological development of the child. N Engl J Med. 341:549–555.
Hajat C. 2011. An introduction to epidemiology. Methods Mol Biol. 713: 27–39.
Hales C, Taylor PN, Channon S, Paradice R, McEwan K, Zhang L, Gyedu M, Bakhsh A, Okosieme O, Muller I, et al. 2018. Controlled antenatal thyroid screening II: effect of treating maternal suboptimal thyroid function on child cognition. J Clin Endocrinol Metab. 103:1583–1591.
Hamed SA. 2015. The effect of antiepileptic drugs on thyroid hormonal function: causes and implications. Expert Rev Clin Pharmacol. 8: 741–750.
Hamilton TE, Davis S, Onstad L, Kopecky KJ. 2008. Thyrotropin levels in a population with no clinical, autoantibody, or ultrasonographic evidence of thyroid disease: implications for the diagnosis of subclinical hypothyroidism. J Clin Endocrinol Metab. 93:1224–1230.
Henrichs J, Bongers-Schokking JJ, Schenk JJ, Ghassabian A, Schmidt HG, Visser TJ, Hooijkaas H, de Muinck Keizer-Schrama SMPF, Hofman A, Jaddoe VVW, et al. 2010. Maternal thyroid function during early pregnancy and cognitive functioning in early childhood: the Generation R Study. J Clin Endocrinol Metab. 95:4227–4234.
Herin F, Boutet-Robinet E, Levant A, Dulaurent S, Manika M, Galatry-Bouju F, Caron P, Soulat JM. 2011. Thyroid function tests in persons with occupational exposure to fipronil. Thyroid. 21:701–706.
Ichihara K, Ozarda Y, Barth JH, Klee G, Qiu L, Erasmus R, Borai A, Evgina S, Ashavaid T, Khan D, et al.; Committee on Reference Intervals and Decision Limits, International Federation of Clinical Chemistry and Laboratory Medicine. 2017. A global multicenter study on reference values: 1. Assessment of methods for derivation and comparison of reference intervals. Clin Chim Acta. 467:70–82.
Isojärvi JI, Pakarinen AJ, Rautio A, Pelkonen O, Myllyla VV. 1995. Serum sex hormone levels after replacing carbamazepine with oxcarbazepine. Eur J Clin Pharmacol. 47:461–464.
Isojärvi JI, Turkka J, Pakarinen AJ, Kotila M, Rättyä J, Myllyla VV. 2001. Thyroid function in men taking carbamazepine, oxcarbazepine, or valproate for epilepsy. Epilepsia. 42:930–934.
Jansen TA, Korevaar TIM, Mulder TA, White T, Muetzel RL, Peeters RP, Tiemeier H. 2019. Maternal thyroid function during pregnancy and child brain morphology: a time window-specific analysis of a prospective cohort. Lancet Diabetes Endocrinol. 7:629–637.
Jonklaas J, Razvi S. 2019. Reference intervals in the diagnosis of thyroid dysfunction: treating patients not numbers. Lancet Diabetes Endocrinol. 7:473–483.
Julvez J, Alvarez-Pedrerol M, Rebagliato M, Murcia M, Forns J, Garcia-Esteban R, Lertxundi N, Espada M, Tardón AR, Galán I, et al. 2013. Thyroxine levels during pregnancy in healthy women and early child neurodevelopment. Epidemiol. 24:150–157.
Korevaar TIM, Medici M, de Rijke YB, Visser W, de Muinck Keizer-Schrama SMPF, Jaddoe VWV, Hofman A, Ross HA, Visser WE, Hooijkaas H, et al. 2013. Ethnic differences in maternal thyroid parameters during pregnancy: the Generation R study. J Clin Endocrinol Metab. 98:3678–3686.
Korevaar TIM, Muetzel R, Medici M, Chaker L, Jaddoe VWV, de Rijke YB, Steegers EAP, Visser TJ, White T, Tiemeier H, et al. 2016a. Association

of maternal thyroid function during early pregnancy with offspring IQ and brain morphology in childhood: a population-based prospective cohort study. Lancet Diabetes Endocrinol. 4:35–43.

Korevaar TI, Taylor PN, Dayan CM, Peeters RP. 2016b. An Invitation to join the Consortium on Thyroid and Pregnancy. Eur Thyroid J. 5:277.

Korevaar TIM, Medici M, Visser TJ, Peeters RP. 2017. Thyroid disease in pregnancy: new insights in diagnosis and clinical management. Nat Rev Endocrinol. 13:610–622.

Lain S, Trumpff C, Grosse SD, Olivieri A, Van Vliet G. 2017. Are lower TSH cutoffs in neonatal screening for congenital hypothyroidism warranted? Eur J Endocrinol. 177:D1–D12.

Lain SJ, Rifas-Shiman SL, Pearce EN, Nassar N, Oken E. 2020. Neonatal thyroxine, maternal thyroid function, and cognition in mid-childhood in a US cohort. Matern Child Health J. 24:503–513.

Larkin JG, McKee PJ, Forrest G, Beastall GH, Park BK, Lowrie JI, Lloyd P, Brodie MJ. 1991. Lack of enzyme induction with oxcarbazepine (600 mg daily) in healthy subjects. Br J Clin Pharmacol. 31:65–71.

Larsen PR, Atkinson AJ, Wellman HN, Goldsmith RE. 1970. The effect of diphenylhydantoin on thyroxine metabolism in man. J Clin Invest. 49: 1266–1279.

Laurberg P, Andersen SL, Hindersson P, Nohr EA, Olsen J. 2016. Dynamics and predictors of serum TSH and fT4 reference limits in early pregnancy: A study within the Danish national birth cohort. J Clin Endocrinol Metab. 101:2484–2492.

Lazarus JH, Bestwick JP, Channon S, Paradice R, Maina A, Rees R, Chiusano E, John R, Guaraldo V, George LM, et al. 2012. Antenatal thyroid screening and childhood cognitive function. N Engl J Med. 366: 493–501.

Lazarus J, Brown RS, Daumerie C, Hubalewska-Dydejczyk A, Negro R, Vaidya B. 2014. 2014 European Thyroid Association Guidelines for the management of subclinical hypothyroidism in pregnancy and in children. Eur Thyroid J. 3:76–94.

Levie D, Korevaar TIM, Bath SC, Dalmau-Bueno A, Murcia M, Espada M, Dineva M, Ibarluzea JM, Sunyer J, Tiemeier H, et al. 2018. Thyroid function in early pregnancy, child IQ, and autistic traits: a meta-analysis of individual participant data. J Clin Endocrinol Metab. 103: 2967–2979.

Levie D, Korevaar TIM, Bath SC, Murcia M, Dineva M, Llop S, Espada M, van Herwaarden AE, de Rijke YB, Ibarluzea JM, et al. 2019a. Association of maternal iodine status with child IQ: a meta-analysis of individual participant data. J Clin Endocrinol Metab. 104:5957–5967.

Levie D, Korevaar TIM, Mulder TA, Bath SC, Dineva M, Lopez-Espinosa MJ, Basterrechea M, Santa-Marina L, Rebagliato M, Sunyer J, et al. 2019b. Maternal thyroid function in early pregnancy and child attention-deficit hyperactivity disorder: an individual-participant meta-analysis. Thyroid. 29:1316–1326.

Li Y, Shan Z, Teng W, Yu X, Li Y, Fan C, Teng X, Guo R, Wang H, Li J, et al. 2010. Abnormalities of maternal thyroid function during pregnancy affect neuropsychological development of their children at 25-30 months. Clin Endocrinol (Oxf). 72:825–829.

Luoma PV, Heikkinen JE, Ylöstalo PR. 1980. Changes in thyroid function tests during phenobarbital treatment in late pregnancy. Int J Gynaecol Obstet. 17:462–464.

Männistö T, Surcel HM, Ruokonen A, Vääräsmäki M, Pouta A, Bloigu A, Järvelin MR, Hartikainen AL, Suvanto E. 2011. Early pregnancy reference intervals of thyroid hormone concentrations in a thyroid antibody-negative pregnant population. Thyroid. 21:291–298.

Maraka S, Mwangi R, McCoy RG, Yao X, Sangaralingham LR, Singh Ospina NM, O'Keeffe DTD, Ycaza AEE, Rodriguez-Gutierrez R, Coddington CC, et al. 2017. Thyroid hormone treatment among pregnant women with subclinical hypothyroidism: US national assessment. BMJ. 356:i6865.

Medici M, Korevaar TIM, Visser WE, Visser TJ, Peeters RP. 2015. Thyroid function in pregnancy: what is normal? Clin Chemistry. 61:704–713.

Methley AM, Campbell S, Chew-Graham C, McNally R, Cheraghi-Sohi S. 2014. PICO, PICOS and SPIDER: a comparison study of specificity and sensitivity in three search tools for qualitative systematic reviews. BMC Health Serv Res. 14:579.

Midgley JEM, Toft AD, Larisch R, Dietrich JW, Hoermann R. 2019. Time for a reassessment of the treatment of hypothyroidism. BMC Endocr Disord. 19:37.

Moleti M, Trimarchi F, Vermiglio F. 2014. Thyroid physiology in pregnancy. Endocr Pract. 20:589–596.

Moog NK, Entringer S, Heim C, Wadhwa PD, Kathmann N, Buss C. 2017. Influence of maternal thyroid hormones during gestation on fetal brain development. Neurosci. 342:68–100.

Mughal BB, Fini JB, Demeneix BA. 2018. Thyroid-disrupting chemicals and brain development: an update. Endocr Connect. 7:R160–R186.

Murcia M, Rebagliato M, Iñiguez C, Lopez-Espinosa MJ, Estarlich M, Plaza B, Barona-Vilar C, Espada M, Vioque J, Ballester F. 2011. Effect of iodine supplementation during pregnancy on infant neurodevelopment at 1 year of age. Am J Epidemiol. 173:804–812.

Murcia M, Espada M, Julvez J, Llop S, Lopez-Espinosa MJ, Vioque J, Basterrechea M, Riaño I, González L, Alvarez-Pedrerol M, et al. 2018. Iodine intake from supplements and diet during pregnancy and child cognitive and motor development: the INMA Mother and Child Cohort Study. J Epidemiol Community Health. 72:216–222.

Nelson SM, Haig C, McConnachie A, Sattar N, Ring SM, Smith GD, Lawlor DA, Lindsay RS. 2018. Maternal thyroid function and child educational attainment: prospective cohort study. BMJ. 360:k452.

Noyes PD, Friedman KP, Browne P, Haselman JT, Gilbert ME, Hornung MW, Barone S Jr, Crofton KM, Laws SC, Stoker TE, et al. 2019. Evaluating chemicals for thyroid disruption: opportunities and challenges with in vitro testing and adverse outcome pathway approaches. Environ Health Perspect. 127:95001.

Ohnhaus EE, Bürgi H, Burger A, Studer H. 1981. The effect of antipyrine, phenobarbitol and rifampicin on thyroid hormone metabolism in man. Eur J Clin Invest. 11:381–387.

Ohnhaus EE, Studer H. 1983. A link between liver microsomal enzyme activity and thyroid hormone metabolism in man. Br J Clin Pharmacol. 15:71–76.

Oken E, Braverman LE, Platek D, Mitchell ML, Lee SL, Pearce EN. 2009. Neonatal thyroxine, maternal thyroid function, and child cognition. J Clin Endocrinol Metab. 94:497–503.

Ornoy A, Weinstein-Fudim L, Ergaz Z. 2016. Genetic syndromes, maternal diseases and antenatal factors associated with autism spectrum disorders (ASD). Front Neurosci. 10:170–121.

Päkkilä F, Männistö T, Hartikainen AL, Ruokonen A, Surcel HM, Bloigu A, Vääräsmäki M, Järvelin MR, Moilanen I, Suvanto E. 2015. Maternal and child's thyroid function and child's intellect and scholastic performance. Thyroid. 25:1363–1374.

Patsalos PN, Zakrzewska JM, Elyas AA. 1990. Dose dependent enzyme induction by oxcarbazepine? Eur J Clin Pharmacol. 39:187–188.

Perucca E, Hedges A, Makki KA, Ruprah M, Wilson JF, Richens A. 1984. A comparative study of the relative enzyme inducing properties of anticonvulsant drugs in epileptic patients. Br J Clin Pharmacol. 18: 401–410.

Pop VJ, Kuijpens JL, van Baar AL, Verkerk G, van Son MM, de Vijlder JJ, Vulsma T, Wiersinga WM, Drexhage HA, Vader HL. 1999. Low maternal free thyroxine concentrations during early pregnancy are associated with impaired psychomotor development in infancy. Clin Endocrinol (Oxf). 50:149–155.

Razvi S, Hostalek U. 2019. Therapeutic challenges in the application of serum thyroid stimulating hormone testing in the management of patients with hypothyroidism on replacement thyroid hormone therapy: a review. Curr Med Res Opin. 15:1–6.

Reinisch JM, Sanders SA, Mortensen EL, Rubin DB. 1995. In utero exposure to phenobarbital and intelligence deficits in adult men. JAMA. 274:1518–1525.

Roche. 2009. Reference intervals for children and adults. Elecsys thyroid tests – TSH, FT4, FT3, T3, T-uptake, FT4-index, anti-TPO, Anti-Tg, Tg. Elecsys systems 1010/2010. Modular analytics E170, cobas e411 and cobas e601 analysers. Mannheim: Roche Diagnostics GmbH.

Shorvon S. 2000. Oxcarbazepine: a review. Seizure: Eur J Epilepsy. 9: 75–79.

Spann MN, Cheslack-Postava K, Brown AS. 2020. The association of serologically documented maternal thyroid conditions during pregnancy with bipolar disorder in offspring. Bipolar Disord. 22:621–628.

Spencer L, Bubner T, Bain E, Middleton P. 2015. Screening and subsequent management for thyroid dysfunction pre-pregnancy and during

pregnancy for improving maternal and infant health. Cochrane Database Syst Rev. 14:CD011263.
Stagnaro-Green A, Abalovich M, Alexander E, Azizi F, Mestman J, Negro R, Nixon A, Pearce EN, Soldin OP, Sullivan S, et al; American Thyroid Association Taskforce on Thyroid Disease During Pregnancy and Postpartum. 2011. Guidelines of the American Thyroid Association for the diagnosis and management of thyroid disease during pregnancy and postpartum. Thyroid. 21:1081–1125.
Stockigt JR. 2001. Free thyroid hormone measurement – a critical appraisal. Assessment Thyroid Function Disease. 30:265–289.
Strolin Benedetti M, Whomsley R, Baltes E, Tonner F. 2005. Alteration of thyroid hormone homeostasis by antiepileptic drugs in humans: involvement of glucuronosyltransferase induction. Eur J Clin Pharmacol. 61:863–872.
Su Q, Zhang S, Hu M, Wang Q, Liu N, Shen H, Zhang Y, Zhang M. 2019. Reference range and sociodemographic characteristics of TSH among reproductive age women in rural China. Biol Trace Elem Res. 189:336–343.
Tacconelli E. 2010. Systematic reviews: CRD's guidance for undertaking reviews in health care. Lancet Inf Dis. 10:229.
Thompson W, Russell G, Baragwanath G, Matthews J, Vaidya B, Thompson-Coon J. 2018. Maternal thyroid hormone insufficiency during pregnancy and risk of neurodevelopmental disorders in offspring: a systematic review and meta-analysis. Clin Endocrinol (Oxf). 88:575–584.
Thorpe-Beeston JG, Nicolaides KH, Felton CV, Butler J, McGregor AM. 1991. Maturation of the secretion of thyroid hormone and thyroid-stimulating hormone in the fetus. N Engl J Med. 324:532–536.
US EPA. 2015 Oct. Neurodevelopmental disorders. America's children and the environment, 3rd ed. [accessed 2020 Jun]. https://www.epa.gov/ace/health-neurodevelopmental-disorders-report-contents
US EPA. 2017. United States Environmental Protection Agency. Technical fact sheet – polybrominated diphenyl ethers (PBDEs); EPA 505-F-17-015; 2017 Nov.
Verrotti A, Scaparrotta A, Cofini M, Chiarelli F, Tiboni GM. 2014. Developmental neurotoxicity and anticonvulsant drugs: a possible link. Reprod Toxicol. 48:72–80.
Visser TJ, Kaptein E, Gijzel AL, de Herder WW, Ebner T, Burchell B. 1993. Glucuronidation of thyroid hormone by human bilirubin and phenol UDP-glucuronyltransferase isoenzymes. FEBS Lett. 324:358–360.
Wang P, Gao J, Zhao S, Guo Y, Wang Z, Qi F. 2016. Maternal thyroxine levels during pregnancy and outcomes of cognitive development in children. Mol Neurobiol. 53:2241–2248.
[WHO] World Health Organisation. 2012. Endocrine disruptors and child health. Possible developmental early effects of endocrine disrupters on child health. https://apps.who.int/iris/bitstream/handle/10665/75342/9789241503761_eng.pdf?sequence=1
WHO IPCS. 2004. Disinfectants and disinfectant by-products. Part IV. Chapter 5 – epidemiological studies. Environmental Health Criteria 216. 277–382.
Wilks MF, Roth N, Aicher L, Faust M, Papadaki P, Marchis A, Calliera M, Ginebreda A, Andres S, Kühne R, et al.; HEROIC consortium. 2015. White paper on the promotion of an integrated risk assessment concept in European regulatory frameworks for chemicals. Sci Total Environ. 521-522:211–218.
Williams FL, Watson J, Ogston SA, Visser TJ, Hume R, Willatts P. 2013. Maternal and umbilical cord levels of T4, FT4, TSH, TPOAb, and TgAb in term infants and neurodevelopmental outcome at 5.5 years. J Clin Endocrinol Metab. 98:829–838.
Wong H, Lehman-McKeeman LD, Grubb MF, Grossman SJ, Bhaskaran VM, Solon EG, Shen HS, Gerson RJ, Car BD, Zhao B, et al. 2005. Increased hepatobiliary clearance of unconjugated thyroxine determines DMP 904-induced alterations in thyroid hormone homeostasis in rats. Toxicol Sci. 84:232–242.
Yang X, Meng Y, Zhang Y, Zhang C, Guo F, Yang S, Ding R, Fan JX. 2019. Thyroid function reference ranges during pregnancy in a large Chinese population and comparison with current guidelines. Chin Med J. 132:505–511.
Zhang YX, Shen CH, Lai QL, Fang GL, Ming WJ, Lu RY, Ding MP. 2016. Effects of antiepileptic drug on thyroid hormones in patients with epilepsy: a meta-analysis. Seizure Eur J Epilepsy. 35:72–79.