Organisation for Economic Co-operation and Development

ENV/JM/MONO(2020)16

Unclassified English - Or. English

24 September 2020

ENVIRONMENT DIRECTORATE
JOINT MEETING OF THE CHEMICALS COMMITTEE AND THE WORKING PARTY ON CHEMICALS, PESTICIDES AND BIOTECHNOLOGY

Cancels & replaces the same document of 23 September 2020

Case Study on use of an Integrated Approach to Testing and Assessment (IATA) and New Approach Methods to Inform a Theoretical Read-Across for Dermal Exposure to Propylparaben from Cosmetics

Series on Testing and Assessment
No. 320

JT03465716

OECD Environment, Health and Safety Publications

Series on Testing and Assessment

No. 320

CASE STUDY ON USE OF AN INTEGRATED APPROACH TO TESTING AND ASSESSMENT (IATA) AND NEW APPROACH METHODS TO INFORM A THEORETICAL READ-ACROSS FOR DERMAL EXPOSURE TO PROPYLPARABEN FROM COSMETICS

IOMC

INTER-ORGANIZATION PROGRAMME FOR THE SOUND MANAGEMENT OF CHEMICALS

A cooperative agreement among FAO, ILO, UNDP, UNEP, UNIDO, UNITAR, WHO, World Bank and OECD

Environment Directorate

ORGANISATION FOR ECONOMIC CO-OPERATION AND DEVELOPMENT

Paris 2020

About the OECD

The Organisation for Economic Co-operation and Development (OECD) is an intergovernmental organisation in which representatives of 37 industrialised countries in North and South America, Europe and the Asia and Pacific region, as well as the European Commission, meet to co-ordinate and harmonise policies, discuss issues of mutual concern, and work together to respond to international problems. Most of the OECD's work is carried out by more than 200 specialised committees and working groups composed of member country delegates. Observers from several countries with special status at the OECD, and from interested international organisations, attend many of the OECD's workshops and other meetings. Committees and working groups are served by the OECD Secretariat, located in Paris, France, which is organised into directorates and divisions.

The Environment, Health and Safety Division publishes free-of-charge documents in twelve different series: **Testing and Assessment; Good Laboratory Practice and Compliance Monitoring; Pesticides; Biocides; Risk Management; Harmonisation of Regulatory Oversight in Biotechnology; Safety of Novel Foods and Feeds; Chemical Accidents; Pollutant Release and Transfer Registers; Emission Scenario Documents; Safety of Manufactured Nanomaterials;** and **Adverse Outcome Pathways**. More information about the Environment, Health and Safety Programme and EHS publications is available on the OECD's World Wide Web site (www.oecd.org/chemicalsafety/).

This publication was developed in the IOMC context. The contents do not necessarily reflect the views or stated policies of individual IOMC Participating Organisations.

The Inter-Organisation Programme for the Sound Management of Chemicals (IOMC) was established in 1995 following recommendations made by the 1992 UN Conference on Environment and Development to strengthen co-operation and increase international co-ordination in the field of chemical safety. The Participating Organisations are FAO, ILO, UNDP, UNEP, UNIDO, UNITAR, WHO, World Bank and OECD. The purpose of the IOMC is to promote co-ordination of the policies and activities pursued by the Participating Organisations, jointly or separately, to achieve the sound management of chemicals in relation to human health and the environment.

This publication is available electronically, at no charge.

Also published in the Series on testing and Assessment link

For this and many other Environment,
Health and Safety publications, consult the OECD's
World Wide Web site www.oecd.org/chemicalsafety/

or contact:

OECD Environment Directorate,
Environment, Health and Safety Division
2, rue André-Pascal
75775 Paris cedex 16
France
Fax : (33-1) 44 30 61 80
E-mail : ehscont@oecd.org

Forward

OECD member countries have been making efforts to expand the use of alternative methods in assessing chemicals. The OECD has been developing guidance documents and tools for the use of alternative methods such as (Q)SAR, chemical categories and Adverse Outcome Pathways (AOPs) as a part of Integrated Approaches for Testing and Assessment (IATA). There is a need for the investigation of the practical applicability of these methods/tools for different aspects of regulatory decision-making, and to build upon case studies and assessment experience across jurisdictions.

The objective of the IATA Case Studies Project is to increase experience with the use of IATA by developing case studies, which constitute examples of predictions that are fit for regulatory use. The aim is to create common understanding of using novel methodologies and the generation of considerations/guidance stemming from these case studies.

This case study was developed by Cosmetics Europe (BIAC) for illustrating practical use of IATA and submitted to the 2019 review cycle of the IATA Case Studies Project. This case study was reviewed by the project team. The document was endorsed at the 4th meeting of the Working Party on Hazard Assessment in June 2020.

The following case study was also reviewed in the project in 2019:

1. CASE STUDY ON THE USE OF INTEGRATED APPROACHES FOR TESTING AND ASSESSMENT FOR SYSTEMIC TOXICITY ARISING FROM COSMETIC EXPOSURE TO CAFFEINE, ENV/JM/MONO(2020)17.
2. CASE STUDY ON THE USE OF INTEGRATED APPROACHES FOR TESTING AND ASSESSMENT FOR 90-DAY RAT ORAL REPEATED-DOSE TOXICITY OF CHLOROBENZENE-RELATED CHEMICALS, ENV/JM/MONO(2020)18.
3. CASE STUDY ON THE USE OF INTEGRATED APPROACHES FOR TESTING AND ASSESSMENT TO INFORM READ-ACROSS OF P-ALKYLPHENOLS: REPEATED-DOSE TOXICITY, ENV/JM/MONO(2020)19.
4. CASE STUDY ON THE USE OF INTEGRATED APPROACHES TO TESTING AND ASSESSMENT FOR PREDICTION OF A 90 DAY REPEATED DOSE TOXICITY STUDY (OECD 408) FOR 2-ETHYLBUTYRIC ACID USING A READ-ACROSS APPROACH FROM OTHER BRANCHED CARBOXYLIC ACIDS, ENV/JM/MONO(2020)20.
5. CASE STUDY ON THE USE OF INTEGRATED APPROACHES TO TESTING AND ASSESSMENT FOR READ-ACROSS BASED FILLING OF DEVELOPMENTAL TOXICITY DATA GAP FOR METHYL HEXANOIC ACID, ENV/JM/MONO(2020)21.
6. CASE STUDY ON THE USE OF INTEGRATED APPROACHES TO TESTING AND ASSESSMENT FOR IDENTIFICATION AND CHARACTERISATION OF PARKINSONIAN HAZARD LIABILITY OF DEGUELIN BY AN AOP-BASED TESTING AND READ ACROSS APPROACH, ENV/JM/MONO(2020)22.

7. CASE STUDY ON THE USE OF INTEGRATED APPROACHES TO TESTING AND ASSESSMENT FOR MITOCHONDRIAL COMPLEX-III-MEDIATED NEUROTOXICITY OF AZOXYSTROBIN - READ-ACROSS TO OTHER STROBILURINS, ENV/JM/MONO(2020)23.

These case studies are illustrative examples, and their publication as OECD monographs does not translate into direct acceptance of the methodologies for regulatory purposes across OECD countries. In addition, these cases studies should not be interpreted as official regulatory decisions made by the authoring member countries.

In addition, a considerations document summarising the learnings and lessons of the review experience of the case studies is published with the case studies:

REPORT ON CONSIDERATIONS FROM CASE STUDIES ON INTEGRATED APPROACHES FOR TESTING AND ASSESSMENT (IATA) -Fifth Review Cycle (2019) -, ENV/JM/MONO(2020)24.

This document is published under the responsibility of the Joint Meeting of the Chemicals Committee and Working Party on Chemicals, Pesticides and Biotechnology.

Synopsis

This case study aims at establishing a proof-of-concept for the value added by New Approach Methodologies (NAMs) in read across: the use of *in silico* information, *in vitro* toxicodynamic and toxicokinetic data, toxicogenomics, and bioactivity data to support the biological similarity of analogues and establish potency trends to inform selection of best source chemical from within a category. This approach, along with consideration of aggregate exposure, is used in an example safety assessment of low toxicity chemicals.

The chemical category under consideration is short linear chain parabens. Parabens are esters of para-hydroxybenzoic acid (pHBA) that are widely used as preservatives in many types of products. These chemicals are used in diverse product sectors including agrochemical, pharmaceutical, food and cosmetics. The read-across category is composed of methylparaben (MP), ethylparaben (EP), propylparaben (PP) and butylparaben (BP), and the safety assessment is for their use in cosmetic products. The goal of the read-across is to fill a theoretical data gap for the reproductive toxicity of PP and demonstrate the safety (i.e. a sufficient margin of exposure or MOE) of PP as used in cosmetics.

Parabens have been shown to lack specific target organ toxicity at very high doses in repeated dose toxicity studies. However, they are active in some uterotrophic assays and have been assigned a conservative point of departure (PoD) by the Scientific Committee for Consumer Safety based on a study with BP. In addition, data have emerged in literature that show parabens exhibit low activity in *in vitro* assays relevant for endocrine activity. These *in vitro* assays suggest the parabens possess very weak activity on some nuclear receptors involved in endocrine homeostasis, many orders of magnitude lower than natural oestrogens/androgens. This case study aims to use NAMs to explore this oestrogenic activity as an approach to inform on biological similarity and relative potency of the category members. This relative potency is then used to adjust the reproductive toxicity point of departure (PoD) derived from BP in order to carry out a theoretical risk assessment for PP based on read-across. This approach is meant to establish a proof of concept for the value that NAMs can add to traditional safety assessment.

As stated, the purpose for this safety assessment is to demonstrate a sufficient MOE exists for PP use in cosmetics with respect to reproductive toxicity potential. The safety assessment was conducted according to the *ab initio* framework/workflow described by Berggren *et al.* (2017) and the OECD (2017). In this workflow, Tier 0 involves consideration of the chemical structure of parabens and the collection of available *in silico*, *in vitro* and *in vivo* data as well as the relevant exposure information from the use of parabens in cosmetic products. These data indicated a similarity in the physicochemical properties of the parabens and an association between chain lengths of the alkoxy groups with the binding propensities towards 16 nuclear receptor structures. Overall, docking simulations indicated a homogenous profile of weak endocrine activity. Structurally, parabens belong to Cramer Class 1, and the level of external exposure to parabens was calculated to be higher than the Threshold of Toxicological Concern (TTC) for Cramer Class 1; therefore, the safety assessment progressed to Tier 1 relying on a read-across approach.

For the read-across in Tier 1, the available *in vivo* reproductive toxicity study data on PP was excluded for demonstration purposes. All four parabens were considered suitable analogues for one another based on multiple parameters (e.g. chemical structure similarity,

expert chemistry review, etc.) and thus formed the read-across category. Based on the remaining *in vivo* data set for the parabens, no target organ was identified up to the highest doses tested in general systemic toxicity studies. Several oestrogenicity studies evidenced low activity for parabens, and an *in vivo* reproductive toxicity study (uterotrophic assay) on BP resulted in weak oestrogenic activity (at 600 mg/kg/day by subcutaneous (SC) injection) which was not observed with the other two analogues (MP and EP). The identified PoD in this study was a NOEL of 2 mg/kg/day, which was considered very conservative. Based primarily on Tanimoto similarity, BP was identified as the closest analogue related to the PP target compound.

In Tier 2, NAMs were considered for this read-across case study by comparing the biological profile of the target with that of the source compounds. Available existing *in vitro* and *in silico* data were collected primarily through the use of a Cosmetics Europe-ToxGPS database called "CE-ToxGPS". Other sources included (but were not limited to) the COSMOS database, the OECD QSAR toolbox, CHEMbl, NCCT dashboard, EDTargetNet and Endocrine Disruptome. In addition, new NAM data were generated including *in vitro* toxicodynamic and toxicokinetic data, toxicogenomics, and bioactivity data. NAMs were used to evaluate similarity in metabolism, potential modes of action (MoA), bioactivity across the short-chain parabens, as well as to establish potency trends across the category (a summary of NAMs and the conclusions from these is shown in Table 1). Methods and data from the NAMs are presented herein, together with a discussion of the key results, uncertainty in the data, and their interpretation. Ultimately, physiologically-based biokinetic modelling (PBBK) is used to estimate internal exposure concentrations from the read-across animal PoD and from the consumer exposure estimates to PP. The internal PoD is then used for comparison to the PP internal exposure adjusted for relative potency based on NAM. Finally, the overall read-across justification and integrated conclusion is provided, as well as a final conclusion on the case study as an example safety assessment informed by NAMs of low toxicity chemicals.

Table 1. Summary of data used in the case study.

Data	Result
Tier 0 data	
TTC	Not applicable since exposure exceeds TTC
Physicochemical properties	Some slight differences but comparable physicochemical properties of the four parabens generally substantiate the category suitability.
In silico alerts and docking simulations	Docking simulations indicate a homogenous profile of weak activity for receptors considered by the Endocrine Disruptome tool. Further substantiation of the category grouping.
Tier 1 data	
In vitro skin penetration in non-viable human skin	Systemic exposure after topical application rapid and equivalent across all category members. No adjustments for the systemically available dose needed to extrapolate findings from toxicity studies from BP to PP.
In vitro metabolism in viable human skin	Extensive first-pass cutaneous metabolism - less than 0.5% of the systemically bioavailable amount in the form of the parent paraben.
Primary human hepatocytes	Primary metabolic pathway for all four parabens in hepatic and skin models is hydrolysis to pHBA. Further metabolism of pHBA is minimal.
Human liver S9 and EpiSkin S9	Rate of ester hydrolysis significantly higher in the liver than in skin. Any parent paraben systemically bioavailable after topical application is rapidly hydrolysed once it reaches the liver.
Human plasma protein binding	All four parabens stable in human plasma, with minimal hydrolysis by esterases therein. Binding to plasma proteins is extensive but varies across parabens.
ToxCast data	MP and EP have lower bioactivity in ToxCast assays than PP and BP. pHBA did not demonstrate any significant activity in the assays. The rank order of potency in oestrogen receptor activity is MP < EP < PP < BP. The relative potency scaling factors for oestrogen receptor activity AC10 median values for PP, EP, and MP were calculated to be 0.37, 0.2 and 0.13, respectively, compared to BP (value of 1).

Tier 2 data	
CALUX	All parabens exhibited oestrogenic but not anti-oestrogenic activity in the absence of rat liver S9. pHBA, was inactive in all cases. The oestrogenic potency increased with chain length but the potency was mostly lower in the presence of rat liver S9. None of the paraben had androgenic activity, while they did show anti-androgenic activity which decreased in the presence of rat liver S9. Importantly, incubations with S9 in all cases decrease bioactivity in the EATS panel and pHBA was essentially devoid of significant biological activity.
Transcriptomics	All parabens elicited changes in the expression of a large number of genes in MFC7 cells. They all up-regulated oestrogen response genes, indicating they share potential MoAs related to endocrine effects. The parabens share a high degree of biological similarity across the category members, with a significant overlap in the affected pathways and the biological activity generally increasing with increasing chain length. The changes in gene expression elicited by MP, EP or BP were also regulated in the same direction (up- or down-regulated) by PP (the target chemical). Fewer genes were affected by pHBA; which were mostly different from those affected by the parabens. These results demonstrate the overall biological similarity and thus the validity of the read-across category. The biological activity associated with the target chemical PP was most similar to that of BP.
Exposure Assessment: deterministic and refined probabilistic consumer exposure	a PBBK model was developed and used to estimate the internal plasma concentrations of MP, PP and BP following whole body exposure in lotion
PBBK modelling	The PoD dose of 2 mg/kg/day BP was administered by SC injection to rats. The C_{max} was 2.1 µ M, the AUC was 3.0 µ mole*h/L and the C_{avg} was 0.13 µ M. The internal exposure estimates for PP were: C_{max} = 0.020 µ M, AUC = 0.340 µ mole*h/L, C_{avg} = 0.014 µ M from the SCCS deterministic consumer exposure estimates; C_{max} = 0.0064 µ M, AUC = 0.011 mmole*h/L, C_{avg} = 0.0046 µ M from the Crème deterministic (worst case) consumer exposure estimates; and C_{max} = 0.00058 µ M, AUC = 0.010 µ mole*h/L, C_{avg} = 0.00042 µ M from the Crème probabilistic (realistic) consumer exposure estimates.

Table of Contents

Tables

Figures

1. Introduction

1.1. Background

Parabens are esters of para-hydroxybenzoic acid (pHBA) and are widely used as preservatives in many types of products from diverse product sectors including agrochemical, pharmaceutical, food and cosmetics where product preservation is essential for safety reasons. The current case study example is based on their use in cosmetics. Parabens can be used in cosmetics as a single chemical entity or in combinations of multiple paraben components in a product formulation. In addition, a consumer may use many cosmetic products daily. Therefore, it is appropriate to consider the total aggregate exposure from all cosmetic products containing parabens (whether alone or in combinations) in a safety assessment. The multiple paraben components under consideration in the current case include the short linear chain parabens: methylparaben (MP), ethylparaben (EP), propylparaben (PP), and butylparaben (BP).

Parabens are generally inactive *in vivo* , but data have emerged in the literature that show weak activity on some nuclear receptors involved in endocrine homeostasis, many orders of magnitude lower than natural oestrogens/androgens. If observed at higher levels, this type of activity could conceivably disrupt hormone homeostasis and be relevant for reproductive toxicity potential. However, low *in vitro* activity levels are such that they may not translate to any effects *in vivo*.

1.2. Problem formulation and purpose

Consumers are dermally exposed to parabens in cosmetic formulations. Therefore, an exposure-based risk assessment must be performed to assure safe use. For the purpose of this case study, a theoretical reproductive toxicity data gap exists for PP.

The current case study aims to use NAM data, particularly based on the weak endocrine activity of parabens, to support a category approach to read-across to fill the theoretical *in vivo* data gap for a reproductive toxicity point of departure for the target chemical PP. The chemical category consists of MP, EP, PP, and BP. PP is the target chemical and MP, EP, and BP are the source chemicals.

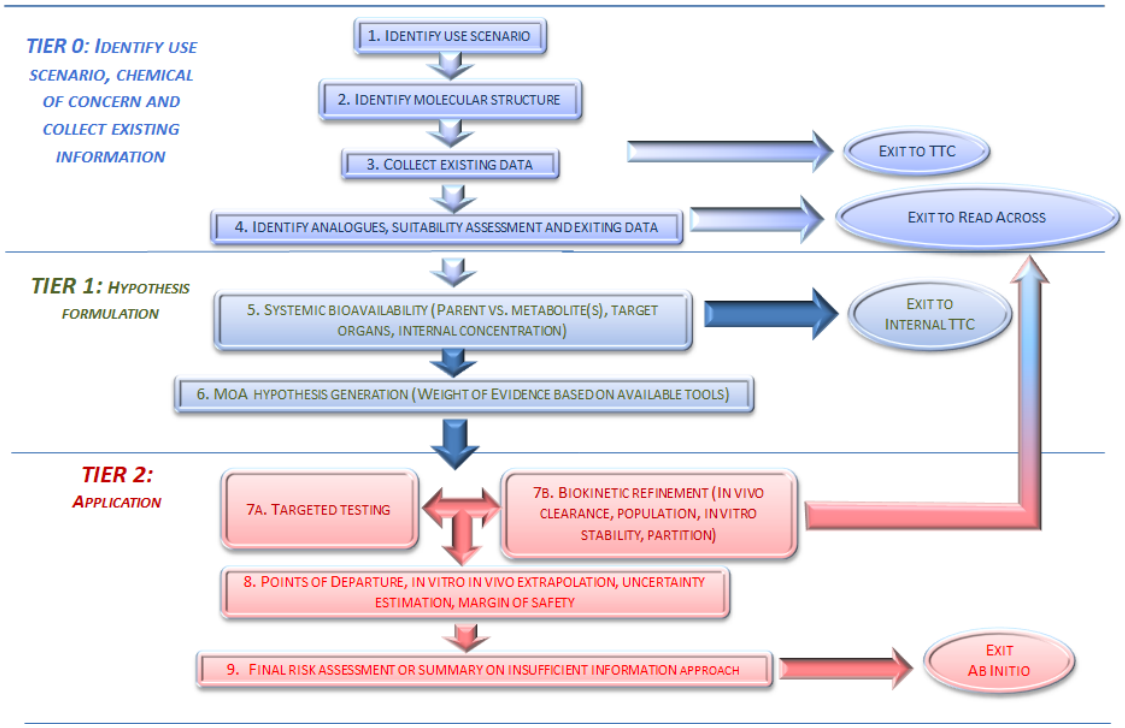
A tiered approach is employed to assess the reproductive toxicity hazard associated with dermal exposure to PP in cosmetics. Tier 0 utilises TTC, and if consumer exposure estimates exceed the TTC level for PP, an appropriate reproductive toxicity point of departure (PoD) is needed for safety assessment. In that case, subsequent tiers will be followed to conduct a read-across informed by NAM. NAMS will also be used in the exposure assessment of parabens. Ultimately, NAM informed internal exposure estimates are compared to calculated margins of internal exposure (MoIE).

The NAM data in this case study is used in the context of research initiatives aimed at developing and implementing non-animal-based tools for safety assessment. The intention is to demonstrate the use of NAMs to support read-across based safety assessment of low toxicity chemicals.

1.3. Use of NAMs in workflow according to OECD Document No. 275

The safety assessment was conducted according to the *ab initio* framework/workflow (Berggren *et al.*, 2017; OECD, 2017) which is organised according to three tiers (Figure 1):

Figure 1. Overview of the *ab initio* framework/workflow for safety assessment.



Tier 0: In this initial tier, a use scenario for the target chemical is defined and existing data on the category are collected. Based on the use scenario, consideration is given to the TTC approach for the safety assessment. The use scenario is conservative and considered as current levels of parabens in all cosmetics products. The resulting exposure is far beyond TTC and therefore the assessment must advance to the next tier and read-across is considered to fill the data gap for reproductive toxicity of PP.

Tier 1: In this tier, absorption, distribution, elimination, and excretion (ADME) and toxicokinetic (TK) and toxicodynamic (TD) properties of the target and source chemicals are considered. ADME/TK data are used to confirm the types and proportions of metabolites formed and to estimate the fraction of remaining parent chemicals after dermal and oral exposure. Existing TD data, including NAMs, are used to evaluate the bioactivity across the category and inform on analogue potency. According to the workflow, in principle a conclusion could be reached at this Tier by comparing internal bioavailability information gained to an internal TTC (iTTC) value (Ellison *et al.*, 2019). However, since iTTC values are not yet established, no conclusion can be reached at this Tier and the assessment must advance to the next tier. But, specific information collected in Tier 1, along with the information gained in Tier 2, plays an essential role in the case study.

Tier 2: In this tier, targeted testing is conducted on TD to (i) verify existing TD data and further inform on biological activity trends shown by previous Tier; and compare estimates

of internal levels of exposure to estimates of safe threshold concentrations adjusted for bioactivity potential.

Specifically, NAMs are used to explore the weak oestrogenic activity of parabens as an approach to inform on relative potency across the category.

This case study example demonstrates the utility of NAMs to inform on several aspects of read-across based safety assessment of low toxicity chemicals. It is also an opportunity to more broadly explore human exposure to parabens and parabens low levels of activity on receptors involved in endocrine homeostasis.

For the purpose of this exercise, no reproductive toxicity data on PP is considered and thus there is a theoretical reproductive toxicity data gap for PP.

For the exposure assessment, dermal exposure is addressed for leave-on cosmetics applied to the whole body. Conservative assumptions are made including that each paraben is present in all cosmetic products at the highest allowable levels. From this conservative assumption, the starting estimate of the aggregate exposure to parabens from combined use of dermally-applied cosmetic products is deterministic and also includes many conservative assumptions regarding product usage. Subsequently refined probabilistic exposure estimates are determined and then potency-adjusted (based on knowledge gained by NAM) in the safety assessment to calculate a MoIE.

1.4. General justification for case study

1.4.1. Endocrine activity of parabens

Parabens are known to possess weak endocrine activity. It is widely understood that reproductive toxicity can occur via many different modes of action (MOA), only some of which include adverse effects on endocrine function and/or hormone homeostasis. In this case study, NAMs are used to explore the weak endocrine activity of short linear chain parabens, but it is acknowledged that a distinction exists between endocrine activity and reproductive toxicity. Herein, the oestrogenic activity of the category members is exploited as a biological proxy of reactivity to investigate potency differences across the category and to apply the information gained from NAMs to the safety assessment.

The potency work in this case study leads to the ranking of MP, EP, and PP against BP, for which a reproductive toxicity PoD is available from legacy *in vivo* data. NAMs will show consistent evidence that BP is the most potent paraben in the category. From NAM data it has been determined that PP (the target chemical for the read-across) has a potency of 0.37 (37%) in comparison to BP (the source chemical for the read-across) with respect to effects related to the oestrogen receptor, which has been shown to be the most sensitive receptor based on dose-response evaluation in NAMs.

1.4.2. Reproductive toxicity of parabens

The reproductive toxicity PoD used in this case study is derived from a data set of reproductive toxicity studies on the category members, primarily uterotrophic assays. The critical study and PoD was selected since it has been used by the Scientific Committee for Consumer Safety (SCCS) as the PoD for risk assessment of parabens in dermally applied cosmetic products. It is a conservative NOEL from a study with BP, and it is applied in read-across to the theoretical reproductive toxicity data gap for PP.

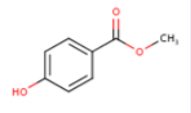
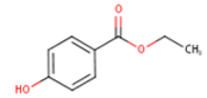
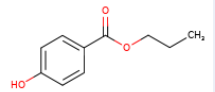
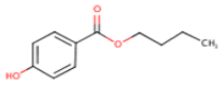
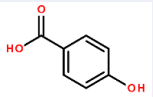
The goal of this case study is primarily to demonstrate how NAMs can be used to support read-across. Specifically, the case study highlights employing PBPK modelling to estimate internal concentrations in both the hazard and exposure assessment and provides an example of the concept of evaluating potency across a category using NAMs. While in the case of short linear chain parabens there is low toxicity and no precise MOA is known, overall this case study addresses the reproductive toxicity apical endpoint through read-across informed by the use of endocrine-oriented NAMs.

2. Tier 0

2.1. Category definition

Parabens are esters of pHBA. The esters link the pHBA moiety to a linear aliphatic alkyl chains of various lengths. The paraben series forming this category are short linear alkyl chain lengths of 1 to 4 carbons. The target and source chemicals are monoconstituent substances (assumed 100% purity) and are structurally identical with the exception of the length of these linear alkyl chains attached via the ester linkage. The target chemical PP has a chain length of 3 carbons. The source chemicals include the two shorter alkyl chain chemicals, MP and EP, and the next longer alkyl chain material (BP) in the series. These source chemicals were identified using a search strategy based on principles described by Wu *et al.* (2010). Tanimoto coefficients, which represent the chemical structural similarity of the source chemicals to the target PP, were derived by comparison of the structural key-based Biovia fingerprints of MP, EP, and BP to PP using a proprietary database of approximately 960 searchable structural keys. The resulting Tanimoto coefficients were 0.81, 0.93, and 0.94, respectively, for MP, EP, and BP (see **Table 2**). The latter, BP, is thus the source chemical presenting the highest structural similarity to PP.

Table 2. Category Members - target and source chemicals.

Target and Source Chemicals	Chemical Name	CAS Number	Molecular Formula	Molecular Weight	Chemical Structure	Similarity (Tanimoto coefficient)
Source 1	Benzoic acid, -4-hydroxy-, methyl ester (Methylparaben; MP)	99-76-3	$C_8H_8O_3$	152		0.81
Source 2	Benzoic acid, -4-hydroxy-, ethyl ester (Ethylparaben; EP)	120-47-8	$C_9H_{10}O_3$	166		0.93
Target	Benzoic acid, -4-hydroxy-, propyl ester (Propylparaben; PP)	94-13-3	$C_{10}H_{12}O_3$	180		1 (target)
Source 3	Benzoic acid, -4-hydroxy-, butyl ester (Butylparaben; BP)	94-26-8	$C_{11}H_{14}O_3$	194		0.94
Common Metabolite	4-Hydroxybenzoic acid (pHBA)	99-96-7	$C_7H_6O_3$	138		Not applicable

2.2. Threshold of Toxicological Concern (TTC)

As a first step in Tier 0, the TTC approach was attempted. The structure of the target chemical PP is evaluated to determine its Cramer Class (Munro *et al.*, 1996). PP is in Cramer Class 1 which has a TTC limit of 0.03 mg/kg/day. The deterministic consumer exposure estimate for PP is calculated in this step. For dermally applied cosmetic products, the typical use scenario is taken which includes 0.19% PP concentration and assumes a

conservative 50% skin penetration, with an applied amount of 17.4 g/day per SCCS Notes of Guidance (SCCS, 2018). This use scenario results in an exposure value of 0.275 mg/kg/day. This exposure estimate is then compared to the TTC limit of 0.03 mg/kg/day. The estimated exposure significantly exceeds the TTC value. Therefore, the TTC approach cannot help conclude on the safety of PP in dermally applied cosmetics and it is necessary to employ read-across. The next step in Tier 0, that must be conducted prior to executing read-across, is to identify suitable analogues for PP and gather the relevant (to reproductive toxicity, the endpoint of interest) existing data on PP and analogues. The category and potential source chemicals have been defined (see section 2.1) and available physico-chemical, *in vivo*, and *in silico* data to inform on analogue similarity are identified below.

2.3. Physicochemical Properties

Physicochemical properties affect bioavailability and consequently biological responses observed *in vitro* or *in vivo*. Therefore, a first step in analysis of proposed analogues for read-across is evaluation of similarity of physicochemical properties (measured and/or calculated). The key physicochemical properties which could affect bioavailability of the four parabens in the category are listed in Table 3.

Table 3. Physicochemical Properties Relevant to Bioavailability.

Parameter	MP	EP	PP	BP
Molecular weight	152.15	166.17	180.2	194.23
Melting Point (°C)	131°C (b)	117°C (b)	97°C (b)	68.5°C
Volatility (mmHg at 25°C)	0.000855 (a)	0.0000929 (a)	0.000307 (a)	0.000251 (a)
Log P	1,96 (a); 1.66 - 1.91 (b)	2.47 (a); 1.81 - 2.57 (b)	3.04 (a), 2.34 - 3.04 (b)	3.57
pKa	8,34 - 8.87 at 25°C (b)	8.18 - 8.9 (b)	7.91 - 8.87 at 25°C (b)	8.34 at 25°C
Aqueous solubility (mg/L)	2500 at 25°C (a); 2335 at 32°C (c)	885 at 25°C (a); 885 (b)	500 at 25°C (a, b); 518 at 32°C (c)	207 at 20°C

a: from EPA EPI Suite; b: from OECD QSAR Toolbox, and from published literature ; c: from

In review of the results for the four parabens (all powders), it can be seen that with increasing side chain length (and thus increasing molecular mass) the logP increases steadily and water solubility decreases. While these properties can impact the relative bioavailability of the parabens, in the case of the target chemical, PP, versus the 3 analogues, the differences are within admitted experimental variation for solubility properties (Dearden & Worth, 2007). The volatility of all four parabens is situated well within the lower end of the semi-volatile class range. The acid/basic behaviour of all four parabens is essentially the same. Since they have a pKa of ~8, although their solubility and thus bioavailability might vary slightly along the GI tract, these short linear chain parabens would be expected to show similar patterns.

2.4. *In vivo* legacy data and rationale for selection of a NOEL for the PoD

A second step in the evaluation is an analysis of similarity of *in silico* alerts that are deemed relevant to the endpoint (in this case reproductive toxicity) being subject to read-across. Key *in vivo* studies on parabens conducted prior to 2017 were reviewed and scored for their reliability per the approach of Klimisch (1997). There are several published regulatory opinions and assessments available for parabens with a focus related to regulatory context available for parabens (reports are available from SCCS, EFSA/FAO, RIVM, etc.); however, these sources of information were not considered here for the sake of simplicity.

As this case study is for demonstration purposes and confined to evaluating potential hazards from dermal exposure to parabens from cosmetics, a minimal relevant *in vivo* data set was selected for use.

Parabens were studied extensively in repeat dose (90 day or equivalent) studies between the 1930s and 1950s. Some of these studies were either not available for review or not reported to sufficient quality to be useful. Some of the more recent studies since Matthews *et al.* (1956) have been included. These studies would not all necessarily meet modern quality criteria, but they provide evidence of a lack of specific target organ toxicity at very high doses. The *in vivo* studies considered are summarised in **Table 4** and **Appendix 6**.

Table 4. *In vivo* data Matrix.

A subset of available Klimisch 1 (K1) and 2 (K2) *in vivo* studies relevant to reproductive toxicity and related endpoints (e.g. repeat dose toxicity) for the paraben category members and their common ester hydrolysis metabolite were reviewed for this case study and are listed here. Some Klimisch 3 (K3) studies were also included.

Test Material	Reference	K Score/Study type and Description	Results	NOAEL (mg/kg/day)	LOAEL (mg/kg/day)
Repeat dose toxicity endpoint					
pHBA	Combined Repeat Dose and Reproductive/Developmental Toxicity Study of pHBA in Rats. OECD SIDS Initial Assessment Report for pHBA (Sponsor Country Japan) (secondary reference). Original reference: Ministry of Health and Welfare (1997).	K1 OECD 422 Combined Repeat Dose and Reproductive/Developmental Toxicity Study. 4-Hydroxybenzoic acid was administered by oral gavage at doses of 30, 200 and 1000 mg/kg/day for 45 days in males and from 14 days before mating until day 3 of lactation in females.	No effects of toxicological significance were observed.	1000	N/A
MP	pHBA esters as preservatives II. Acute and Chronic Toxicity in Dogs, Rats and Mice (Matthews *et al.*, 1956).	K3 A series of studies of MP, EP, PP and BP were conducted in Wistar rats for up to 96 weeks at doses up to 2% and 8% in diet. (Dog and mice data not presented here)	No effects at 2%. At 8% MP, decreased body weight gain and mortality was observed.	NR	NR
EP	pHBA esters as preservatives II. Acute and Chronic Toxicity in Dogs, Rats and Mice (Matthews *et al.*, 1956).	K3 A series of studies of MP, EP, PP and BP were conducted in Wistar rats for up to 96 weeks at doses up to 2% and 8% in diet. (Dog and mice data not presented here)	No effects at 2%. At 8% EP, decreased body weight gain and mortality was observed.	NR	NR
PP	pHBA esters as preservatives II. Acute and Chronic Toxicity in Dogs, Rats and Mice (Matthews *et al.*, 1956).	K3 A series of studies of MP, EP, PP and BP were conducted in Wistar rats for up to 96 weeks at doses up to 2% and 8% in diet. (Dog and mice data not presented here)	No effects at 2%. At 8% PP, decreased body weight gain and mortality was observed.	NR	NR
BP	pHBA esters as preservatives II. Acute and Chronic Toxicity in Dogs, Rats and Mice (Matthews *et al.*, 1956).	K3 A series of studies of MP, EP, PP and BP were conducted in Wistar rats for up to 96 weeks at doses up to 2% and 8% in diet. (Dog and mice data not presented here)	No effects at 2%.	NR	NR
Developmental toxicity endpoint					
BP	Developmental Toxicity Evaluation of BP in Sprague Dawley Rats (Daston, 2004).	K1 OECD 414 Prenatal Developmental Toxicity Study in Sprague Dawley rats at up to 1000 mg/kg/day.	Decreased maternal weight gain at highest dose tested. No differences in developmental parameters.	Maternal = 100 Foetal = 1000	Maternal = 1000
Reproductive toxicity endpoint					
MP	Lack of effect of MP on the reproductive system in male rats (Hoberman *et al.*, 2008).	K1 Non-guideline study. Male rats fed diets containing 10000 ppm test material from day 22 of age for 56	No effects observed on reproductive organs or parameters.	10000 ppm in the diet (equivalent to	N/A

		days. Weekly measurement of serum LH, FSH and testosterone. After 56 days animals were sacrificed, sex organs were weighed and evaluated by histopathology including tubular staging of testis. Sperm evaluations were conducted including concentration and motility, daily sperm production, and morphology.		~1088 mg/kg/day)	
BP	Lack of effect of BP on the reproductive system in male rats (Hoberman *et al.*, 2008).	K1 Non-guideline study. Male rats fed diets containing 10000 ppm test material from day 22 of age for 56 days. Weekly measurement of serum LH, FSH and testosterone. After 56 days animals were sacrificed, sex organs were weighed and evaluated by histopathology including tubular staging of testis. Sperm evaluations were conducted including concentration and motility, daily sperm production, and morphology.	No effects observed on reproductive organs or parameters.	10000 ppm in the diet (equivalent to ~1088 mg/kg/day)	N/A
MP	Lack of oestrogenic effects of food preservatives (parabens) in uterotrophic assays (Hossaini *et al.*, 2000).	K2 OECD 440 Uterotrophic Study. MP and PP tested up to 100 mg/kg oral and sub cutaneously (SC). EP tested up to 100 mg/kg SC and 1000mg/kg oral.	No effects of MP at any dose tested	100 mk/Kg/day	
EP	Lack of oestrogenic effects of food preservatives (parabens) in uterotrophic assays (Hossaini *et al.*, 2000).	K2 OECD 440 Uterotrophic Study.). EP tested up to 100 mg/kg SC and 1000mg/kg oral.	No effects of EP at any dose tested.	100 mg/Kg/day	
PP	Lack of oestrogenic effects of food preservatives (parabens) in uterotrophic assays (Hossaini *et al.*, 2000).	K2 OECD 440 Uterotrophic Study. PP tested up to 100 mg/kg oral and sub cutaneously (SC).	No effects of PP at any dose tested.	100 mg/Kg/day	
BP	Lack of oestrogenic effects of food preservatives (parabens) in uterotrophic assays (Hossaini *et al.*, 2000).	K2 OECD 440 Uterotrophic Study. BP tested up to 600mg/kg SC. (Combinations also tested but not reported here.)	BP showed a weak oestrogenic effect at 600 mg/kg/day by SC injection.	400 mg/Kg/day (SC) for BP	600 mg/Kg/day (SC) for BP
MP	*In vivo* oestrogen bioactivities of alkyl parabens (Lemini *et al.*, 2003).	K2 Similar to OECD 440 except use of immature rats and mice and ovariectomised mice. Study of MP administered by SC injection to immature or ovariectomised CD1 mice and immature Wistar rats. (Results for ovariectomised mice or immature rats reported here.)	Weak oestrogenic activity of parabens administered by SC injection was observed. P-HBA was inactive in immature rats.	Lowest NOEL was: 5.5 mg/Kg/day for MP	LOELs ranging from 16.5 – 165 mg/Kg/d for MP
EP	*In vivo* oestrogen bioactivities of alkyl parabens (Lemini *et al.*, 2003).	K2 Similar to OECD 440 except use of immature rats and mice and ovariectomised mice. Study of EP administered by SC injection to immature or ovariectomised CD1 mice and immature Wistar rats. (Results for ovariectomised mice or immature rats reported here.)	Weak oestrogenic activity of parabens administered by SC injection was observed. P-HBA was inactive in immature rats.	Lowest NOEL was: 0.6 mg/Kg/day for EP	LOELs ranging from 6 -180 mg/Kg/day for EP
PP	*In vivo* oestrogen bioactivities of alkyl parabens (Lemini *et al.*, 2003).	K2 Similar to OECD 440 except use of immature rats and mice and ovariectomised mice. Study of PP administered by SC injection to immature or ovariectomised CD1 mice and immature Wistar rats. (Results for ovariectomised mice or immature rats reported here.)	Weak oestrogenic activity of parabens administered by SC injection was observed. P-HBA was inactive in immature rats.	Lowest NOEL was: 6.5mg/Kg/day for PP	LOELs ranging from 20-65mg/Kg/day for PP

BP	*In vivo* oestrogen bioactivities of alkyl parabens (Lemini *et al.*, 2003).	K2 Similar to OECD 440 except use of immature rats and mice and ovariectomised mice. Study of BP administered by SC injection to immature or ovariectomised CD1 mice and immature Wistar rats. (Results for ovariectomised mice or immature rats reported here.)	Weak oestrogenic activity of parabens administered by SC injection was observed. P-HBA was inactive in immature rats.	Lowest NOEL was: 7 mg/Kg/day BP	LOELs ranging from 7 – 70 mg/Kg/day for BP
p-HBA	*In vivo* oestrogen bioactivities of alkyl parabens (Lemini *et al.*, 2003).	K2 Similar to OECD 440 except use of immature rats and mice and ovariectomised mice. Study of MP, EP, PP and BP administered by SC injection to immature or ovariectomised CD1 mice and immature Wistar rats. (Results for ovariectomised mice or immature rats reported here.)	Weak oestrogenic activity of parabens administered by SC injection was observed. P-HBA was inactive in immature rats.	N/A	N/A
MP	Morphometric analysis of mice uteri treated with the preservatives MP, EP, PP and BP (Lemini *et al.*, 2004).	K2 OECD 440 Uterotrophic Study. MP administered by SC injection (SC) to ovarectomised CD1 mice.	Weak oestrogenic activity of SC administered parabens was observed.	N/A	LOEL was 55 mg/Kg/day for MP.
EP	Morphometric analysis of mice uteri treated with the preservatives MP, EP, PP and BP (Lemini *et al.*, 2004).	K2 OECD 440 Uterotrophic Study. EP administered by SC injection (SC) to ovarectomised CD1 mice.	Weak oestrogenic activity of SC administered parabens was observed.	N/A	LOEL was 60 mg/Kg/day for EP
PP	Morphometric analysis of mice uteri treated with the preservatives MP, EP, PP and BP (Lemini *et al.*, 2004).	K2 OECD 440 Uterotrophic Study. PP administered by SC injection (SC) to ovarectomised CD1 mice.	Weak oestrogenic activity of SC administered parabens was observed.	N/A	LOEL was 65 mg/Kg/day for PP
BP	Morphometric analysis of mice uteri treated with the preservatives MP, EP, PP and BP (Lemini *et al.*, 2004).	K2 OECD 440 Uterotrophic Study. BP administered by SC injection (SC) to ovarectomised CD1 mice.	Weak oestrogenic activity of SC administered parabens was observed.	N/A	LOEL was 70 mg/Kg/day for BP
MP	Some alkyl hydroxybenzoate preservatives (parabens) are oestrogenic (Routledge *et al.*, 1998).	K2 OECD 440 Uterotrophic Study. MP administered orally (up to 800 mg/kg/day) and by SC injection (up to 1200mg/kg/day) to immature female Alpk:AP rats.	Weak oestrogenic activity of parabens was observed.	NOEL for MP 800 mg/kg/day oral route	N/A for MP
BP	Some alkyl hydroxybenzoate preservatives (parabens) are oestrogenic (Routledge *et al.*, 1998).	K2 OECD 440 Uterotrophic Study. BP administered orally (up to 1200mg/kg/day) and by SC injection (up to 1200mg/kg/day) to immature female Alpk:AP rats.	Weak oestrogenic activity of parabens was observed.	NOEL for BP 800 mg/Kg/day (oral) and 40 mg/Kg/day (SC)	BP 1200 mg/Kg/day (oral)
pHBA	Oestrogenic effects of pHBA in CD1 Mice (Lemini *et al.*, 1997).	K2 Similar to OECD 440 Uterotrophic Study except use of immature mice.	Weak activity observed; inconsistent with *in vitro* data for pHBA.	N/A	50
BP	Effect of neonatal exposure to oestrogenic compounds on development of excurrent ducts of the rat testis through puberty to adulthood (Fisher *et al.*, 1999).	K3 Non guideline study. Neonatal Wistar rats were administered by SC injection with 2mg/kg/day BP in corn oil on PNDs 2-18. Animals were sacrificed on day 18 and the testes and epididymides removed. Testis weights were recorded. AQP-1 immunoexpression was measured and excurrent duct morphology examined.	NR	2 mg/kg/day was the NOEL in this study adopted by SCCS as conservative PoD for risk assessment of BP and PP (SCCS, 2013)	NR

NR = Not reported, N/A = Not Applicable.

The majority of the *in vivo* legacy reproductive toxicity data presented in **Table 4**, **Appendix 6** correspond to Level 4 assays according to the OECD Conceptual Framework on Endocrine Disrupters (OECD, 2012). While the *in vivo* data generally demonstrate low toxicity, a number of the reproductive toxicity studies evidenced activity for some of the

parabens. For example, a study on BP (Hossaini *et al.*, 2000) exhibited a weak oestrogenic activity at 600 mg/kg/day by SC injection which was not seen with the other analogues. According to another study (Hoberman *et al.*, 2008) no reproductive toxicity could be evidenced for either MP or BP, but in this study testosterone and luteinizing hormone were statistically significantly reduced at the highest concentration tested while thyroid hormones and follicle stimulating hormone were statistically increased compared to control group. However, none of these hormonal changes resulted in adversities in the animals.

The NOEL of 2 mg/kg/day for BP from the single low dose study by Fisher *et al.* (1999) was chosen by the Scientific Committee for Consumer Safety (SCCS, 2013) as a very conservative PoD for risk assessment of parabens, while others have used values up to 1000 mg/kg/day as a PoD for BP (CIR, 2017). The SCCS acknowledged that the choice of this PoD was very conservative and unusual in terms of the SCCS Notes of Guidance (SCCS, 2018) and general principles of risk assessment, but it was chosen in the absence of more appropriate data at the time. For the purpose of this case study demonstration, this limited *in vivo* data set and the SCCS conservative PoD for risk assessment was chosen and it was considered sufficient. In addition, the results of these *in vivo* studies prompted the consideration of NAMs related to endocrine activity (particularly EATS) for this read-across case study in order to compare the biological profiles and potencies of the target compound with those of the source compounds.

2.5. *In silico* alerts

In addition to evaluation of physicochemical properties and available *in vivo* data, an analysis of *in silico* data that are deemed relevant to the endpoint that is the subject of the read-across is important in determining the similarity and suitability of identified analogues. In this case study, the *in silico* alerts related to the reported weak endocrine properties observed in the *in vivo* data for these parabens were evaluated. The profilers that the OECD QSAR Toolbox highlights as pertinent for reproductive toxicity – i.e. the DART scheme, Estrogen Receptor, Binding, Retinoic Acid Receptor Binding – and the rtER Expert System from U.S. EPA were evaluated to examine the similarity among the category members. The results of key *in silico* profiling of the four parabens in the category and their common ester hydrolysis metabolite pHBA are listed in **Table 5**. Not surprisingly, the parabens exhibited binding propensities for the oestrogen receptor; however, they were outside the applicability domain of the RAR-profiler. The ER profilers showed that the short linear chain parabens displayed a trend across putative affinities for the oestrogen receptor as a function of alkyl chain length. pHBA was an outlier with respect to the parabens. While binding predictions by these ER profilers do not necessarily translate directly into *in vivo* apical adverse effects, the predictions add to the weight of evidence that the category grouping is pertinent. The similarity of the ER binding responses across the analogues support BP as the most appropriate source substance (from a biological similarity viewpoint) for PP.

Table 5. *In silico* Profilers Relevant to Reproductive Toxicity.

Profiling results obtained from OECD QSAR Toolbox v 4.2

Chemical	DART scheme	Estrogen Receptor Binding	Retinoic Acid Receptor Binding	rtER Expert System USEPA
pHBA	Not known precedent reproductive and developmental toxic potential	Weak binder, OH group	Not possible to classify according to these rules	No alert found
MP	Known precedent reproductive and developmental toxic potential >> 4-alkylphenol-like derivatives (2b-3)	Weak binder, OH group	Not possible to classify according to these rules	Parabens
EP	Known precedent reproductive and developmental toxic potential >> 4-alkylphenol-like derivatives (2b-3)	Weak binder, OH group	Not possible to classify according to these rules	Parabens
PP	Known precedent reproductive and developmental toxic potential >> 4-alkylphenol-like derivatives (2b-3)	Moderate binder, OH group	Not possible to classify according to these rules	Parabens
BP	Known precedent reproductive and developmental toxic potential >> 4-alkylphenol-like derivatives (2b-3)	Moderate binder, OH group	Not possible to classify according to these rules	Parabens

Profilers & (applicability to endpoint): DART scheme (Developmental and Reproductive Toxicity), Estrogen Receptor Binding (Toxicity to Reproduction), Retinoic Acid Receptor Binding (Toxicity to Reproduction) rtER expert system (Toxicity to Reproduction)

To further explore oestrogen receptor binding propensities of the parabens, docking simulations were performed using the online docking tool 'Endocrine Disruptome' (http://endocrinedisruptome.ki.si/). The Endocrine Disruptome provides predictions of binding probabilities as a function of atomic-level information that is extracted from the three-dimensional structures of the ligand and the included nuclear receptors (Kolšek, 2014). Therefore, Endocrine Disruptome has a very large applicability domain while providing semi-quantitative predictions. These properties, together with the possibility of inspecting docked poses, makes it a more insightful tool than other QSAR models that usually simply discriminate between binders and nonbinders. These docking simulations were used to characterise the binding propensities of short linear chain parabens and their common ester hydrolysis metabolite pHBA towards the sixteen structures, belonging to twelve nuclear receptors. The structure of the chemical was drawn using the graphical interface of the tool and then submitted to docking simulations.

Docking simulations were repeated five times for each chemical and a visual inspection of the docked poses highlighted plausible binding modes. Docking scores are a sum of intermolecular and intramolecular contributions within the ligand binding pocket and the underlying algorithm attempts to identify the global minimum of such a sum (Trott and Olson, 2010). The key-assumption of any virtual docking approach is that docking scores are effective in discriminating binders (low docking scores) from non-binders (high docking scores). More precisely, the Endocrine Disruptome tool established three thresholds for the AutoDock docking scores that enables the classification of binding propensities into four probability classes (Kolšek *et al.*, 2014). These thresholds were established according to a conservative approach since Kolšek and co-authors decided that the true-positive rate was a more important than the true-negative rate for the division of the probability classes. The arithmetic mean of the five docking scores was retained as the final score for the quantitative description of the binding affinities of chemicals. These final scores were then compared to critical score thresholds (specific for each receptor) and associated with color-coded binding probability classes: green, yellow, orange and red.

These colours indicate low, low intermediate, high intermediate and high binding probabilities, respectively.

The docking simulation results are in **Table 6**. They show that all four parabens are associated with a low binding probability class (green colour). The only exception to this tendency is represented by the androgen receptor (AR) in antagonistic conformation (AR an.) which is associated with the lower class of intermediate binding probability (yellow colour). (AR activity was originally surveyed based on the hormonal changes observed in the study by Hoberman *et al.* (2008) and the alerts gathered by the *in silico* approaches. Those results were investigated further separately but for simplicity of this case study demonstration the results of those investigations are not included.) An increase in chain length of the alkoxy groups is associated with decreasing docking scores, reflective of greater binding. The decrease in docking scores across the homologue series of chemicals indicates an increasing gain in favourable hydrophobic interactions between parabens and the binding pockets of the receptors as the length of the alkoxy moieties increases.

Table 6. Docking scores towards sixteen structures belonging to twelve nuclear receptors for pHBA and short chain parabens.

	pHBA	MP	EP	PP	BP
AR	-6.0	Ethyl- -6.1	-6.3	-6.6	-6.8
AR an	**-5.9**	**-5.9**	**-6.0**	**-6.3**	**-6.3**
ER α	-5.6	-5.7	-6.0	-6.5	-6.7
ER α an.	-5.6	-5.7	-6.0	-6.4	-6.5
ER β	-5.7	-5.8	-6.1	-6.4	-6.5
ER β an.	-5.6	-5.7	-6.1	-6.4	-6.5
GR	-5.7	-5.9	-6.1	-6.3	-6.4
GR an.	-5.2	-5.4	-5.6	-5.8	-5.8
LXR α	-5.5	-5.6	-5.9	-6.3	-6.4
LXR β	-6.0	-6.1	-6.3	-6.7	-6.9
PPAR α	-5.5	-5.5	-5.8	-6.4	-6.5
PPAR β	-5.8	-5.7	-6.0	-6.0	-6.0
PPAR γ	-5.3	-5.4	-5.9	-6.5	-6.7
RXR α	-6.3	-6.0	-6.5	-6.9	-7.0
TR α	-5.9	-5.9	-6.3	-6.7	-6.8
TR β	-5.7	-5.8	-6.1	-6.6	-6.7

Docking simulations performed using the online docking tool 'Endocrine Disruptome' (http://endocrinedisruptome.ki.si/).
Green and yellow indicate low and intermediate binding probabilities respectively. The code "an." indicates receptors in antagonistic conformations.
AR = androgen receptor; ER = oestrogen receptor; GR = glucocorticoid receptor; LXR = Liver X receptor; PPAR = peroxisome proliferator-activated receptor; RXR = retinoid X receptor; TR = thyroid hormone receptor

2.6. Summary of the outcome of Tier 0

The following conclusions can be made from Tier 0:

- TTC: The estimated exposure exceeds the TTC. Since the TTC approach cannot help conclude on safety, the next step in the assessment approach is to consider read-across.

- Physicochemical properties: While there are some slight differences, comparison of the physicochemical properties across the four parabens generally substantiates the suitability of the category.
- *In vivo* data: While there was no target organ identified up to the highest doses tested in general systemic toxicity studies, there is some evidence of weak oestrogenic activity for MP, EP, and BP in several uterotrophic studies. A conservative NOEL of 2 mg/kg/day is established based on a study where effects were observed after administration by SC injection (thus by-passing first pass metabolism) of 600 mg/kg/day of BP. However, in another study BP did not exhibit reproductive toxicity. In terms of Tanimoto similarity, BP was identified as the closest compound to the target PP.
- *In silico* alerts and docking simulations: Overall, docking simulations indicate a homogenous profile of weak activity with for the receptors considered by the Endocrine Disruptome tool, further substantiating the suitability of the four parabens to form one category.

3. Tier 1

3.1. Read-across hypothesis

The starting hypothesis for this category is that, based on their highly similar chemical structure, the target chemical PP will have similar bioavailability and bioactivity as the source chemicals MP, EP and BP. The key aspects of the hypothesis are as follows:

- The similar chemical structure and physicochemical characteristics will result in similar bioavailability, metabolism and reactivity, which results in similar biological and functional effects.
- The available *in vivo* systemic toxicity data generally demonstrate the similar biological activity across the category. The same or similar MoA across category members are responsible for the observed effects.
- The parent category members are metabolised by ester hydrolysis via endogenous esterases in the skin or systemically after absorption, with all four parabens producing a common primary metabolite, pHBA, and similar corresponding short linear chain alcohols.
- The rate and extent of ester hydrolysis will be similar across parabens, resulting in similar exposures to the common metabolite pHBA, which does not contribute significantly to the observed toxicity.
- Chain length differences in the parent esters and in the other primary metabolite (i.e. the linear aliphatic alcohol) across the parabens will result in a predictable potency trend in observed effects across category members with increasing alkyl chain length.

Available scientific evidence from traditional *in vivo* data and NAMs addressing these aspects of the read-across hypothesis will provide a weight of evidence that data from BP can be read across to fill the theoretical reproductive toxicity data gap for PP. Further, NAM data will provide evidence of a potency trend across the parabens, and that relative potency information can be used to inform the read-across based safety assessment.

3.2. Comparison of ADME properties of parabens

Understanding absorption, distribution, metabolism and excretion (ADME) properties and the relative rate and extent of biotransformation across the parabens is important in the examination of potential potency differences across the category members that could arise from differences in bioavailability and internal exposure levels. In this case study, the ADME data generated from NAMs is used to compare the behaviour of the four parabens in the category using similar *in vitro* conditions and to provide information helpful in the selection of the most appropriate source chemical for the read-across.

The metabolism of parabens in humans is well-studied (Abbas *et al.*, 2010; Moos *et al.*, 2016). It is understood that parabens undergo ester hydrolysis to form a common primary metabolite, pHBA, and a corresponding linear aliphatic alcohol. This understanding was the starting point for the *in vitro* ADME and toxicokinetics (TK) evaluations performed to compare these characteristics across the paraben category in context of this case study.

A PBBK model for MP, EP and BP was previously developed by Campbell *et al.* (2015) and applied in this case study. Available (in the literature) skin penetration parameters from Cross and Roberts (2000) were used rather than values from *in vitro* ADME data generated from NAMs presented in this case study. (In this case study, the *in vitro* ADME data is solely used to compare the behaviour of the four parabens in the category using similar *in vitro* conditions and to provide information helpful in the selection of the most appropriate source chemical for the read-across).

3.2.1. Metabolism in skin

The following experimental data for the parabens were considered:

(1) *ex vivo* absorption through human skin, and

(2) biotransformation *in vitro*, in human skin explants, primary human hepatocytes (PHH), and S9 from 3D human skin cultures and liver.

Results from penetration studies using non-viable human skin (Appendix 1, Figure 3 and Table 10) indicate the rapid penetration of MP and PP into and through skin is rapid which was comparable when applied in the same aqueous buffered saline solvent. This suggests that systemically available amounts of these short-chain parabens are similar and that no adjustments regarding dermal penetration properties are needed. In addition, if more volatile solvents than water are used (e.g. ethanol), the dermal delivery may be decreased.

To better understand dermal bioavailability, potential first-pass metabolism in the skin should be considered as well as the dermal penetration (Manwaring *et al.*, 2015). Experiments in which PP was applied to viable human skin explants and incubated for 24 h showed that it is extensively metabolised by cutaneous enzymes, such that, after 24 h, nearly all PP applied to human skin was subsequently present in the medium as metabolites (Appendix 1, Table 11). Twelve metabolites were detected in the culture medium, several of which were identified by high-resolution mass spectrometry (Appendix 1, Table 12). The major metabolite identified was pHBA, which accounted for 75% of the total metabolites in the medium and 42% of the applied dose. The findings from the *ex vivo* metabolism studies were in accordance with those from short-term (4 hour) incubations with S9 from the reconstructed human epidermal skin model, EpiSkin S9. In the latter, each paraben was readily metabolised, with each paraben exhibiting a similar initial rate of depletion (Appendix 1, Figure 4), with the production of pHBA. These *ex vivo* and *in vitro* studies indicate that (i) very low amounts of parent paraben enter the systemic circulation after topical application due to first-pass metabolism in the skin, and (ii) the major metabolite entering the systemic circulation is pHBA.

It should be noted here that these *in vitro* ADME data in skin are used solely to compare the metabolic behaviour of the four parabens in the category using similar *in vitro* conditions. Subsequent PBBK modelling performed for the parabens (see Section 4.) utilised skin penetration parameters that were available in the literature (Cross and Roberts, 2000).

3.2.2. Metabolism in hepatocytes and liver S9

Once entering the systemic circulation, a compound can be further metabolised by the liver. The *in vitro* intrinsic clearance was also examined for the parabens since this is an important determinant in liver metabolism. In two *in vitro* hepatic models, primary human hepatocytes (PHH) and human liver S9, the intrinsic clearance ($CL_{int,\ in\ vitro}$) values indicated that all four parabens are high hepatic clearance compounds (Appendix 1, Table 13).

As with metabolism in *ex vivo* skin, in incubations with PHH, pHBA accounted for the majority of metabolite formed, since depletion of parent was concomitant with an increase in pHBA formation (Appendix 1, Figure 5). These findings are in accordance with others who have studied the metabolism of several parabens in human liver microsomes (Abbas *et al.*, 2010). This finding is also in line with the substrate specificity of the liver carboxylesterase-1 (CES1) isoform, which has been reported to prefer esters with a large, bulky acyl group and a small alcohol group (Laizure *et al.*, 2013).

3.2.3. Metabolism comparison in liver and skin S9

The metabolism of parabens in liver S9 and EpiSkin S9 were compared under the same incubation conditions. These incubations were undertaken as a screening assay to provide an indication of the metabolic stability of the chemicals in liver and skin. In addition, this assay provided some comparative information on the xenobiotic metabolizing that were responsible for paraben metabolism in liver- and skin-based models (Eilstein *et al.*, 2020). The rate of metabolism of the four parabens was much higher (between 70- and 210-fold higher) in liver than in EpiSkin S9 (Appendix 1, Table 13). The reason for the lower rate of metabolism of short linear chain parabens in EpiSkin S9 compared to the liver S9 may be attributed to the carboxylesterase isoform, carboxylesterases-2 (CES2), known to be mainly expressed in the skin (Fagerberg *et al.*, 2014). CES2 prefers lipophilic substrates with a large alcohol group (Laizure *et al.*, 2013; Taketani *et al.*, 2007) rather than small alcohol groups as are present in the parabens. It is likely the metabolism of the parabens in EpiSkin S9 is mediated by CES2.

There were two metabolites common to PHH, liver S9 and EpiSkin S9, namely pHBA and a direct sulphate conjugate of the parabens S9 (Appendix 1, Table 14). Oxidation (most likely of the alkyl-chain of the molecule rather on the ring moiety (Moos *et al.*, 2016)), was evident in incubations with PHH and liver S9 but not in EpiSkin S9. This may be expected considering the much lower abundance and activities of CYP enzymes in skin compared to the liver (Hewitt *et al.*, 2013). There was some indication that pHBA was further metabolised in incubations with EP and BP (Appendix 1, Figure 5), possibly to the glycine, glucuronide and/or sulphate conjugates, which are known minor human urinary metabolites (Brand *et al.*, 2018). However, when pHBA was incubated with PHH, it was not metabolised and no conjugates were detected (Appendix 1, Table 14). This finding again indicates that pHBA is the major metabolite and that further metabolites are minor.

3.2.4. Metabolism in plasma

In addition to undergoing metabolism in the liver, some esters are hydrolysed by esterases in the plasma (Fu *et al.*, 2016). Since human plasma is reported not to contain carboxylesterases (Li *et al.*, 2005), the parabens may be substrates for other esterases known to be present e.g. butyrylcholinesterase, paraoxonase, and albumin esterase (Li *et al.*, 2005). However, when incubated with human plasma, all four parabens were stable in plasma and less than 6% of the parabens were hydrolysed to pHBA (Appendix 1, Table 14, stability controls). The degree of plasma protein binding was high and increased with increasing paraben chain length (Appendix 1, Table 15). The fraction bound suggests that the free fraction *in vivo* could vary between 26% for MP to only 4% for BP. The observed binding of pHBA to plasma proteins was much lower than that of any of the parent chemicals.

3.3. Bioactivity in ToxCast (potential MoA of parabens)

To explore biological activity and survey potential MoAs, efforts to find biological data for PP and similar chemicals were undertaken using ToxCast (U.S. EPA). ToxCast data was of particular interest also to increase confidence in the similarities of the structurally related chemicals in the category. As PP was not tested in all ToxCast assays this approach cannot be considered to afford a comprehensive biological coverage; nonetheless, results from 656 assays can give some meaningful insights into mode of action.

Initially, a structure similarity search utilising Accelrys Isentris was employed to identify molecules similar to PP, the target for read-across, that also had ToxCast data. Specifically, the structurally similar compounds were defined on the basis of 960 specific structural features, whereby the degree of structural similarity depends on the number of searchable keys that a stored structure has in common with the query, compared to the total number of searchable keys. For the purpose of this exercise a similarity cut-off of >50 keys was used. A total of 24 chemicals were identified with eight of these being parabens of varying chain lengths including the three source chemicals in the category (identified earlier by Tanimoto and expert chemical review) and the common metabolite pHBA.

Analysis of the ToxCast data associated with MP, EP, BP and pHBA was undertaken. The analysis focused on assay hits with no flags as reported by the U.S. EPA. Flags associated with response data from ToxCast assays indicates potential issues with the fit model (false positive/false negative) as identified by the U.S. EPA and this can result in significant uncertainties in the interpretation of the data. Therefore, for this case study, only response data without flags were included. No data on pHBA was included as there were no hits without flags, and the results for the parabens are listed in Table 7.

Table 7. ToxCast hit counts for parabens

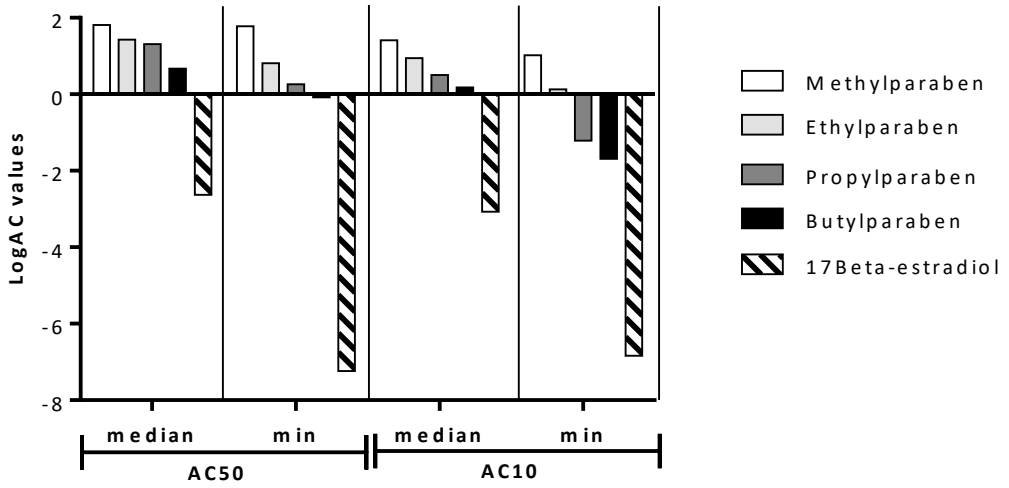
Name	CAS	ToxCast Chemical ID	Similarity Cutoff (Isentris)	Result Count	Hits without flags	% hits relative to all assays
PP	94-13-3	22527	100	656	31	4.73
BP	94-26-8	20209	>80	1357	95	7.00
EP	120-47-8	22528	>70	1279	38	2.97
MP	99-76-3	22529	>60	783	9	1.15

Based on the percentage of hits relative to total number of assays in which the compounds were tested, MP (1.15%) and EP (2.97%) appear to have lower bioactivity in ToxCast assays than PP (4.73%) and BP (7.00%).

Next the assay hits across the parabens were compared, which showed that commonality was consistently observed in relation to the oestrogen receptor activity (data not shown). Due to this convergence, the ToxCast oestrogen receptor model was explored further through collaboration with the U.S. EPA. Full details of the oestrogen receptor model are described elsewhere (Browne *et al.*, 2015). Briefly, the results from 18 oestrogen receptor ToxCast high-throughput screening assays, measuring different points along the signalling pathway with different assay technologies, are integrated into a computational model to discriminate chemicals on the basis of their relative oestrogen receptor bioactivity. For this analysis of the parabens, the resulting oestrogen receptor activity is shown alongside a known oestrogen receptor agonist, 17beta-estradiol, for comparison. Results (see Figure 2) demonstrate that the rank order of potency for oestrogen receptor activity is MP < EP < PP < BP, with the reference substance showing, as expected, much greater oestrogen receptor activity overall.

Figure 2. Relative oestrogen receptor bioactivity in ToxCast.

AC values are shown in the table below the figure.



	MP	EP	PP	BP	17Beta-estradiol
AC50.median	1.81	1.43	1.31	0.67	-2.63
AC50.min	1.78	0.81	0.26	-0.08	-7.24
AC10.median	1.41	0.95	0.50	0.18	-3.07
AC10.min	1.02	0.13	-1.21	-1.69	-6.84

The AC10 and AC50 values listed in Figure 2 were derived by R. Judson (personal communication). BP is associated with the lowest concentrations for both AC10 and AC50 in comparison to the other parabens. Thus, it is assigned a potency of 1 relative to the other category members (see Table 8).

Table 8. Calculation of potency scaling factors from ToxCast oestrogen receptor activity data

	AC10.median	Calculated Scaling (potency) Factor*
BP	0.184926581	1
PP	0.503476501	0.37
EP	0.946787935	0.20
MP	1.405220807	0.13

*calculated as 1/(AC10 of individual paraben/AC10 of BP)

ToxCast data traditionally rely on the concentration of chemical associated with 50% of maximum activity, i.e. AC50. However, since this assay response reflects agonist effects on the oestrogen receptor, increasing concentrations trigger increasing oestrogen activity. The AC10 relates to concentrations associated with 10% of maximum activity or effect on oestrogen receptor, and they are lower concentrations than those at the AC50. Therefore, using the AC10 value is more conservative in the case of a risk assessment and it can be considered more protective. As a result, the AC10 median data were selected as the basis for the potency comparisons of the parabens. Relative to BP, which is the most potent in

the category and assigned a scaling factor of 1, PP is assigned a scaling factor of 0.37, followed by EP and MP, with scaling factors of 0.2 and 0.13, respectively.

These calculated relative potency scaling factors are employed later in the case study for the subsequent safety assessment.

3.4. Summary of the outcome of Tier 1

The following conclusions can be made from Tier 1:

- Read-across: the ADME properties of the parabens support the similarity hypothesis of the category.
- *In vitro* ADME:
  - Systemic exposure to parabens after topical application in aqueous solvent is rapid and at similar finite doses, it is equivalent across all category members. This suggests no adjustments for the systemically available dose are needed to extrapolate findings from toxicity studies from BP to PP.
  - Extensive first-pass cutaneous metabolism results in less than 0.5% of the systemically bioavailable amount to be in the form of the parent paraben.
  - The primary metabolic pathway for all four parabens in hepatic and skin models is hydrolysis to pHBA, confirming existing literature. Further metabolism of pHBA is minimal.
  - The rate of ester hydrolysis is significantly higher in the liver than in skin, indicating that any parent paraben that is systemically bioavailable after topical application is rapidly hydrolysed once it reaches the liver.
  - All four parabens are stable in human plasma, with minimal hydrolysis by esterases therein. Binding to plasma proteins is extensive; however, the free fraction *in vivo* could vary across the paraben category members, which should be taken into account in predicting *in vivo* clearance values.
- ToxCast:
  - MP and EP appear to have lower bioactivity in ToxCast assays than PP and BP, and pHBA did not demonstrate any significant activity in the assays.
  - The rank order of potency in oestrogen receptor activity in the ToxCast high-throughput screening assays is MP < EP < PP < BP.
  - Based on AC10 median values in the ToxCast oestrogen receptor activity assays, BP was assigned a scaling factor of 1 as the most potent of these parabens. The relative potency scaling factors for the AC10 median values for PP, EP, and MP were calculated at 0.37, 0.2 and 0.13, respectively, compared to BP.

4. Tier 2

4.1. Targeted testing in the CALUX assays

In investigating potential MoAs for reproductive toxicity, an obvious consideration is steroid hormones and their receptors, particularly the androgen and oestrogen receptors. These receptors can be modulated in their activity by synthetic chemicals and other xenobiotics, as well as by endogenous molecules. Based on this notion, and the oestrogen activity of the parabens observed in the molecular docking and bioactivity (ToxCast) data already gathered, specific CALUX® transactivation assays were selected to examine the similarities and differences in the endocrine activity of parabens. Use of CALUX® transactivation assays has been shown to be able to reduce the need of the much larger ToxCast panels of less specific assays to measure similar endpoints (Collet *et al.*, 2019; van Vugt-Lussenburg *et al.*, 2018). Since endocrine activities represent molecular initiating events rather than more downstream key events in some reproductive toxicity adverse outcome pathways (AOPs), evaluating endocrine activity is a way to survey many potential AOPs simultaneously (van der Burg *et al.*, 2015). As such, interaction with receptors for oestrogen-, androgen-, thyroid signalling and steroidogenesis (EATS) are relevant to potential MOAs based on endocrine activity. A range of CALUX assays, complemented with specific assays to measure thyroid and steroidogenesis interferences, was selected to create a complete EATS panel in which the parabens were evaluated (see Appendix 2 for assay details and methods). The outcomes of the CALUX assays are listed below.

4.1.1. Cytotoxicity

Data on cytotoxicity was included (as it is the case for all assays) in order to consider biologically relevant concentrations, investigate the toxicities of parabens and explore whether the primary metabolite, pHBA has a role in the biological effects elicited by the parabens.

In the cytotoxicity CALUX assay (Appendix 2, Figure 6, Table 16), toxicity was only observed for the two longest chain parabens at concentrations $>10^{-4}$ M. In the presence of rat liver S9 the cytotoxicity decreased, indicating that the metabolites are less cytotoxic than the parent compounds. This is supported by the fact that their main metabolite, pHBA, shows no cytotoxicity on the cytotox CALUX up to $1x10^{-3}$ M.

4.1.2. Oestrogen and androgen activity

All parabens showed oestrogenic activity in the absence of rat liver S9, but no anti-oestrogenic activity was observed (Appendix 2, Figure 7). The oestrogenic potency increased with chain length; most compounds had a PC10 value in the lower- or sub-micromolar range. In the presence of a metabolic fraction, however, most parabens were metabolised into less potent oestrogens. The metabolite, pHBA, was inactive in all cases.

The AR CALUX assay showed that none of the compounds had androgenic activity, while they did show anti-androgenic activity (Appendix 2, Figure 8). The observed activity was in the lower micromolar range for all parabens, but not for the metabolite, pHBA. Similar to that observed for oestrogenic activity, the anti-androgenic activity also decreased in the presence of rat liver S9.

4.1.3. Thyroidogenic activity

Given that some steroidogenic activity was observed in MP and BP compared to the control group, the biological similarity related to Thyroidogenic activity were investigated.

No significant thyroidogenic activity was detected for any of the compounds, and anti-thyroidogenic activity was observed for MP (Appendix 2, Table 16). Also, for the second thyroid-related assay, hTPO inhibition, little activity was observed (Appendix 2, Table 16). MP showed a 20% decrease in signal only at the highest tested concentration. Inhibition of T4 binding to transthyretin (TTR) was observed for all four parabens. The potency of all compounds was similar, with PC20 values in micromolar range. Only MP was 10- to 100-fold less potent. The metabolite pHBA did not show any activity on the thyroid hormone receptor β (TRβ) and TTR binding assays, but TPO inhibition was observed for this compound at high concentration.

4.1.4. Steroidogenic activity

All four parabens affected steroidogenesis following exposure of H295R cells and subsequent quantification of 17beta-estradiol and/or testosterone production using the oestrogen receptor-α and androgen receptor CALUX bioassay. The effect most often observed was an increase in the oestrogen production. EP and PP additionally decreased the production of androgens. However, according to OECD guidelines, two consecutive active concentrations are required to identify a compound as 'positive'; using this definition, none of these parabens significantly decreased testosterone production, and only MP, EP and PP significantly increased oestrogen production. The metabolite, pHBA, resulted in marginally increased oestrogen production at the highest tested concentration, and as such would also score 'negative'.

4.1.5. EATS conclusions

Importantly, incubations with S9 in all cases decrease bioactivity in the EATS panel (Appendix 2, Table 16). This is consistent with the fact that the major metabolite, pHBA, is devoid of significant biological activity and only shows slight activity in the TPO- and H295R assay at millimolar concentrations. Conversely, the four parabens tested were shown to be hormonally active *in vitro*, acting as oestrogens and anti-androgens. Little direct effect on thyroid receptor signalling and hTPO inhibition was observed but TTR binding was found positive and the parabens were able to influence steroid production according to the H295R assay.

The parent parabens all exhibited measurable activity as agonists in the oestrogen receptor assay when tested at high concentrations (Appendix 2, Figure 7 and Table 16), while being antagonists in the AR assay at high concentrations (Appendix 2, Figure 8 and Table 16). This linked activity has been noted before in other endocrine active substances (Sohoni & Sumpter, 1998). While it can be argued that anti-androgenic activity in some cases may contribute to the oestrogenicity of a substance *in vivo*, in the case of the short linear chain parabens the anti-androgenic activity observed in the EATS panel is of comparatively low potency relative to the observed oestrogenic activity. Both the oestrogenic and anti-androgenic effect of the parabens decreased significantly in the presence of rat liver S9, suggesting that the parabens are readily metabolised to inactive metabolites. The EATS results generally demonstrate that endocrine activity increases with increasing chain length, suggesting a trend in potency across the category.

The results from the EATS assays supported the earlier findings of ER activity from *in silico* alerts and ToxCast data. Based on these data, in subsequent toxicogenomics analyses conducted to further evaluate the toxicodynamic properties of the individual parabens, the focus remained on gaining more understanding of the similarities and differences in the endocrine (particularly oestrogenic) activity across the category members in addition to surveying their biological activity more broadly.

4.2. Transcriptional profiling

A transcriptomics approach was used to inform on the broad biological activity and similarity of the parent parabens, as well as of the common metabolite pHBA. In the toxicogenomics work there was also interest in examining the ability of the category members to activate the oestrogen receptor as demonstrated in the ToxCast and CALUX assay data and implied in the *in silico* evaluations. In the toxicogenomic experiments, the transcriptional response of MCF7 cells was evaluated after exposure to each of the parabens or pHBA (see Appendix 3). The use of MCF7 cells offers a reasonable *in vitro* system to assess the broad biological activity of the parabens as well as further explore their endocrine activity potential since these cells express multiple nuclear hormone receptors as well as other regulatory proteins.

The results of the toxicogenomics analyses indicate that each of the parabens is able to elicit changes in the expression of a large number of genes, as compared to controls, particularly at the highest dose tested. With regard to establishing biological similarity, the transcriptional profile elicited by each of the parabens shares a high degree of similarity across the category members.

A significant number of genes whose expression is up-or down-regulated by MP, EP or BP is also regulated in the same direction (up- or down-regulated) by PP (the target chemical). This is shown in the Eisen diagram heat map (Appendix 3, Figure 9) of the genes (up-regulated in red; down-regulated in blue) whose expression was modified in the MCF7 cells exposed to the indicated parabens (at the highest doses tested) for 6 h. In the case of comparing the gene changes elicited by pHBA to those elicited by the parent parabens, there are clearly fewer genes affected by pHBA (Appendix 3, Figure 9). Figure 9 also demonstrates that there are increased gene changes in MCF7 cells across the parabens as the chain length increases. This is a general indication that the biological activity of the short linear chain parabens increases with increasing chain length.

Comparing the toxicogenomic data across the 4 parabens, there are 133 common genes identified whose expression is modified by each of the parabens in a significant manner and in the same direction (66 genes up-regulated and 67 genes down-regulated) (data not shown). In order to more closely examine the similarities of the differentially expressed genes between the potential source chemicals for the read across and the target chemical, a one to one comparison of the transcriptional profiles of each source paraben (MP, EP, and BP) was made against the transcriptional profile of PP. When compared to PP, MP elicited changes in the expression of 360 common genes, EP elicited changes in expression of 256 common genes, and BP elicited changes in expression of 634 common genes (data not shown). The results indicate highest numbers of commonly affected genes were between BP and PP, where 319 genes were up-regulated and 315 genes were down-regulated.

The main metabolite of these parabens, pHBA, also elicited significant gene expression changes at the highest concentration evaluated (615 genes total, 312 were up-regulated and 303 down-regulated). However, the gene expression changes from pHBA are mostly

different than the ones elicited by any of the parabens. Comparing the transcriptional profile pHBA with that of each of the parabens, the expression of only 45 genes was modified in the same direction (19 up-regulated and 26 down-regulated), although at a different magnitude (data not shown).

To determine the most important biological activities (based on these gene changes) of each of the parabens in the category, the transcriptional profile identified for each of the parabens was analysed for pathway enrichment. For simplicity only the top six enriched gene sets identified with only the up-regulated genes are shown (Appendix 3, Table 16, Table 17, Table 18 and Table 19). Looking broadly across the four parabens, there is significant overlap in the affected pathways, indicative of their overall biological similarity and thus the validity of the read-across category. The top Hallmark pathways that are most up-regulated by the parabens are: oestrogen response early and late, TNFA signalling via NFKB, unfolded protein response, hypoxia, androgen response, glycolysis, epithelial mesenchymal transition, IL2 STAT5 signalling, and MTORC1 signalling. Once again, the biological activity profile of BP has the highest similarity to that of PP (Appendix 3, Table 19 and Table 20).

Comparison of the results from the gene expression and pathway analyses demonstrates that similar transcriptomic responses are elicited by exposure to the parent parabens. The toxicogenomics data provides evidence of strong concordance in the biological activity of the category members as identified by transcriptional profiling of MCF7 cells exposed to the parabens. In addition, the transcriptomic profiles of the parabens clearly demonstrate they share an ability to up-regulate oestrogen response genes in MCF7 cells. These transcriptomic results support the read-across category hypothesis with regard to broad biological similarity, and more specifically provide evidence that the parabens share potential MoAs related to endocrine effects. Further, in one-to-one comparisons between the target chemical PP and the potential source parabens MP, EP, and BP, the biological activity associated with the target chemical PP is most similar to that of BP. This information aligns with the high structural similarity of these parabens (based on Tanimoto) and supports the read-across hypothesis that it is appropriate to use data from BP in read-across to PP. In conclusion, these toxicogenomics data provide additional evidence of biological similarity across the category.

4.3. Read-across conclusion informed by NAM

Available toxicokinetic and toxicodynamic information presented herein on the short linear chain parabens in this category support that these materials are biologically similar in addition to being structurally similar. NAM data have provided useful information to facilitate the selection of BP as the most appropriate analogue in the category to be the source chemical in the read across to the PP target chemical. The NAM data provide a weight of evidence to support the key aspects of the read-across hypothesis as described earlier:

- The similar chemical structure across the category members leads to similarities in physicochemical characteristics, metabolism, and reactivity which results in similar biological and functional effects. *The data supporting this conclusion includes physicochemical predictions and measurements, results of in silico alerts and receptor docking predictions, and absorption, distribution, metabolism, and excretion (ADME) data from the literature for human in vivo metabolism and from NAM in vitro assays.*

- The available *in vivo* systemic toxicity data generally demonstrate the similar biological activity across the category. *Supporting evidence is provided by the in vivo animal data from the literature.*
- The parent category members (MP, EP, PP and BP) are similarly metabolised by ester hydrolysis via endogenous esterases in the skin or systemically after absorption to produce the same common primary metabolite, pHBA. *The data supporting this includes ADME information from the literature and NAM in vitro assay results in primary human hepatocytes and liver S9 and skin S9 experiments.*
- The primary metabolite, pHBA, does not contribute to the observed toxicity for parabens. *The data supporting this includes ADME information from the literature and NAM data demonstrating trends of decreasing metabolism and increasing potency with increasing chain length across the category members.*
- The parabens exhibit a predictable potency trend in observed effects across category members with increasing alkyl chain length. *The quantitative trend observed in ToxCast data was confirmed by EATS (CALUX) assays results and qualitatively substantiated by multiple types of NAM data.*
- Category members are expected to exhibit the same or similar biological activity and MoAs responsible for observed effects. *This conclusion is supported by in silico data, in vitro NAM data exploring ADME and bioactivity, and Toxicogenomics.*

These conclusions support that it is appropriately conservative to read across data from the source chemical BP to fill a data gap for the target chemical PP.

For the purpose of the safety assessment in this case study, a reproductive toxicity PoD of 2 mg/kg/day will be used for PP based on read-across from BP. This PoD is used as a starting point in PBBK modelling to estimate internal exposures in the rat following SC injection administration of 2 mg/kg/day of BP.

4.4. Exposure assessment for dermally applied cosmetics

An exposure assessment was performed for consumer exposure from dermally applied cosmetics. Complete details can be found in Appendix 4. Briefly, the assessment includes both deterministic consumer exposure values for each of the parabens as cited by SCCS (2018), as well as refine probabilistic values derived from the Crème model.

For the purpose of the safety assessment in this case study, both exposure estimates (deterministic and probabilistic) will be used. These external exposure values are used as a starting point in the PBBK modelling to estimate internal exposures in humans following dermal exposure to PP.

4.5. PBBK modeling in exposure scenarios

Physiologically-based biokinetic (PBBK) models are mathematical models used to quantify the absorption, distribution, metabolism and excretion of a chemical inside the body following exposure. They are constructed as an interconnected system of compartments representing various tissues described by mass balance differential equations that are solved to predict the amount of chemical in each compartment over time (Gerlowski & Jain, 1983). The physiological basis of this modelling approach allows

internal concentrations resulting from external exposures to be predicted, allowing comparisons including across species and exposure routes.

The physiological structure of PBBK models provides a particularly useful framework for conducting cross species extrapolations (Clewell & Andersen, 1985). The application of PBBK models to support interspecies extrapolation depends on the concept of target tissue exposure equivalence; that is, in the absence of pharmacodynamic (susceptibility) differences, the toxicity of a chemical in different species is expected to be associated with similar concentrations of the chemical (or its toxic metabolite) in the tissue where the toxicity is observed (Clewell *et al.*, 2002). In cases of general systemic toxicity, or where the target tissue has not been identified, the concentration in the blood can be used to represent the target tissue exposure. While acute effects may depend on the maximum concentration achieved in the tissue, longer-term toxicity is generally associated with the average concentration over time, which can be calculated as the area under the curve (AUC) divided by the duration of the exposure. The toxic mode of action determines whether the concentration of interest is that of the parent chemical, a stable metabolite, or a reactive metabolite (Clewell, 2005). To apply a PBBK model for interspecies extrapolation, the model is first used to simulate the exposure of interest (dose, route, and duration) in the experimental species, and the internal dose metric (peak or average concentration) is calculated. The parameters in the PBBK model are then changed to those for the target species of concern and the dose is adjusted until the same internal dose metric is achieved. The dose that produces the same internal dose metric is then considered the kinetically equivalent dose (Clewell *et al.*, 2002).

The details of the PBBK model applied in this case study to estimate internal concentrations of parabens resulting from external (applied) exposures in humans (from dermally applied cosmetics) and rat (from subcutaneous injection) are provided in Appendix 5.

4.5.1. PBBK modelling in consumer exposure to parabens from dermally applied cosmetics

In the case of dermal exposure to parabens from cosmetic products, a PBBK model was developed and used to estimate the internal plasma concentrations of MP, PP and BP following whole body exposure in lotion (Appendix 5, Figure 13). A body surface area of 13845 cm^2 was used to simulate the whole body of a 60 kg human. The MP exposure was estimated assuming 17.4 g paraben-containing product use per day, and a concentration of 0.44% MP, 0.48% EP, 0.18% PP, or 0.19% BP. This amounts to a twice daily application of 2.8 $\mu g/cm^2$ MP, 3.0 $\mu g/cm^2$ EP, 1.1 $\mu g/cm^2$ PP, or 1.2 $\mu g/cm^2$ BP exposures. Due to different kinetics simulated for MP, the time to reach a steady peak concentration for consecutive days was longer for MP compared to the other parabens. The information in Appendix 5, Figure 13 is summarised in Appendix 5, Table 29 by the area under the curve (AUC) and average concentration (C_{avg}) on the last day of exposure, and the maximum concentration (C_{max}). Exposure estimates generated using the Crème Global exposure model were also used as input to the PBBK model. The results of these simulations are shown in Appendix 5, Table 30. For PP, the internal exposure estimates were C_{max} of 0.020 µ M, the AUC was 0.340 µ mole*h/L and the C_{avg} was 0.014 µ M from the SCCS deterministic consumer exposure estimates; C_{max} of 0.0064 µ M, the AUC was 0.110 mmole*h/L and the C_{avg} was 0.0046 µ M from the Crème deterministic (worst case) consumer exposure estimates; and C_{max} of 0.00058 µ M, the AUC was 0.010 µ mole*h/L and the Cavg was 0.00042 µ M from the Crème probabilistic (realistic) consumer exposure estimates.

4.5.2. PBBK modelling in rats after subcutaneous exposure to parabens

Based on read-across from BP, the conservative PoD of 2.0 mg/kg/day can be used for risk assessment of reproductive toxicity potential for PP. The results of simulating the exposure scenario in the rat toxicity study identifying the BP NOEL of 2.0 mg/kg/day (Fisher *et al.*, 1999) is shown in Appendix 5, Figure 14. The dose of 2 mg/kg/day BP was administered by SC injection in rats. The simulation results show the plasma time-course curve and summary pharmacokinetic parameters. From these, the C_{max} was 2.1 µ M, the AUC was 3.0 µ mole*h/L and the C_{avg} was 0.13 µ M.

4.6. Calculating margins of internal exposure (MoIE)

4.6.1. Metrics used in traditional risk assessment for calculating MoE

Traditional risk assessments, using animal data, rely on the comparison of the human exposure to either a point of departure (PoD) in an animal study or a reference dose (RfD) that is derived from the PoD, after application of various established uncertainty factors (UFs) to account for areas of uncertainty in extrapolation of the data. For cosmetic safety evaluation according to SCCS (2018), the risk assessment is typically based on calculation of a Margin of Exposure (or Safety) (MoE or MoS). The MoE is defined by the PoD divided by the human exposure (PoD/Exposure). A safe risk ratio (RR) is generally considered to be a MoE ≥ 100 when the PoD is determined in an animal study. These traditional metrics for risk assessment rely on nominal/external doses of chemicals, both at the level of the PoD (e.g. administered dose in the animal dose-response study) and of the exposure (e.g. estimated or measured human exposure to the chemical).

4.6.2. Metrics used with PBBK informed risk assessment for calculating MoIE

In this case study, PBBK modelling was used to estimate blood concentrations following external exposures to BP in experimental animals and PP in humans. From these internal exposure estimates, it is possible to calculate a Margin of Internal Exposure (MoIE). A MoIE differs from a traditional margin of exposure (MoE) in that it is calculated as the ratio of a measure of internal exposure, such as blood concentration or target-tissue dose, rather than a measure of external exposure concentration or ingested dose (Bessems *et al.* 2017). The ability to rely on a measure of internal rather than external exposure reduces the uncertainty in the risk assessment by incorporating chemical-specific information on the uptake, distribution, metabolism and excretion of the chemical in both the experimental animal and the human (Clewell & Clewell, 2008). In particular, calculation of internal exposures with a PBBK model can be used to replace the default uncertainty factor of 4 for interspecies differences in TK differences (IPCS 2005, WHO 2010). The U.S. EPA follows this practice in determining Reference Concentrations and Reference Doses (U.S. EPA 2011). Thus, a MoIE of 25 would be equivalent to the default external dose MoE of 100, but with added benefit of greater precision for the specific chemical of concern.

4.6.3. Calculated margins of internal exposure (MoIE) – safety assessment for PP in the case study

From the PBBK modelling, it has been concluded that external SC exposure in rats to 2 mg/kg/day of BP (the NOEL) results in an internal exposure of 2.1 µ M. Similarly, using PBBK modelling, the human exposure simulation suggests an internal exposure of 0.020 µ M to the target chemical PP (see Appendix 5, Table 29) when using the SCCS deterministic values. When using the refined probabilistic consumer exposure evaluation for the realistic

exposure scenario (i.e. scenario 2d), the human exposure simulation suggests an internal exposure of 0.0064 µ M for conservative exposure assumptions (scenario 2a) and 0.00058 µ M for realistic exposure assumptions (scenario 2d) (see Appendix 5, Table 30).

Based on the relative potency information on the parabens that was gained in the NAM evaluations, the internal exposure can further be adjusted for potency as appropriate, prior to calculating the risk ratio. The relative potency trends observed in multiple NAM data sets supported that the biological activity of the parabens is similar but increases with increasing alkyl chain length. This was particularly demonstrated based on NAM evaluations of the weak endocrine activity of parabens, in particular in ER activity evaluated in ToxCast and CALUX assays. As the risk assessment endpoint is reproductive toxicity, endocrine activity is a relevant potential MoA. The ERalpha CALUX assay has been validated (Sonneveld *et al.*, 2005, 2006, Besselink, 2015) and is included in OECD TG455 for both the agonistic and antagonistic mode, with a dose range finding run and two subsequent runs (OECD, 2016, Annex 4). This test can also be combined with S9 metabolism (Van Vugt-Lussenburg *et al.*, 2018). For these reasons it was employed in the context of a larger study with multiple chemicals, with single runs for all chemicals for reasons of resource efficiency, thereby deviating from OECD TG455. Since the generated data are reliable as the CV of this assay is below 5% and accuracy is 95-100% for agonism and antagonism, they have been used in the weight of evidence to support the potency trend among target and source compounds and positive control. Nevertheless, since ToxCast data was generated in duplicate runs, as a conservative approach it was used to establish potency across the related parabens. Therefore, the relative ER bioactivity based on ToxCast AC10 values (see Table 7) is used as a basis for the potency adjustment. The scaling potency factor for the target chemical PP as compared to the source chemical contributing the animal PoD, BP, is 0.37. Taking this approach, the following MoIE calculation is used:

$$\text{MoIE} = C_{max}\text{rat BP}/[(C_{max}\text{human PP x (Relative Potency of PP/BP)}]$$

Following a deterministic consumer exposure estimate, the internal MoIE is calculated:

Type of exposure assessment: SCCS, Tier 1 deterministic			
PoD	Internal exposure	Relative Potency	MoIE
C_{max} rat for BP: 2.1 µM	C_{max} human PP: 2.0E-2µM	Factor for PP: 0.37	MoIE = 2.1/(2.0E-2*0.37) = 284

Following probabilistic consumer exposure estimates for worst case and realistic scenarios, the internal MoIE is calculated as:

Type of exposure assessment: Crème model, Tier 2a probabilistic (worst case scenario)			
PoD	Internal exposure	Relative Potency	MoIE
C_{max} rat for BP: 2.1 µM	C_{max} human PP: 6.4E-3 µM	Factor for PP: 0.37	MoIE = 2.1 / (6.4E-3*0.37) = 887
Type of exposure assessment: Crème model, Tier 2d probabilistic (realistic scenario)			
PoD	Internal exposure	Relative Potency	MoIE
C_{max} rat for BP: 2.1 µM	C_{max} human PP: 5.8E-4 µM	Factor for PP: 0.37	MoIE = 2.1 / (5.8E-4*0.37) = 9786

The resulting MoIEs are 284 using the SCCS deterministic values (see Appendix 4, Table 22), and using the Crème model probabilistic Tier 2a and 2d consumer exposure estimates (see Appendix 4, Table 23) the MoIEs are 887 and 9786, respectively.

Based on the safety assessment conducted in this case study, the MoIE would be considered sufficient to support the use of PP in dermally applied cosmetic products.

4.7. Summary of the outcome of Tier 2

The following conclusions can be made from Tier 2:

CALUX:

- All 4 parabens were of low toxicity, with observed decreases in viability at concentrations >10^{-4} M. pHBA, was not cytotoxic up to 1x10^{-3} M and incubations with rat liver S9 decreased the cytotoxicity, indicating that metabolites are less cytotoxic than the parent parabens.
- All parabens exhibited oestrogenic but not anti-oestrogenic activity in the absence of rat liver S9. pHBA, was inactive in all cases.
- The oestrogenic potency increased with chain length but the potency was mostly lower in the presence of rat liver S9.
- None of the compounds had androgenic activity, while they did show anti-androgenic activity which decreased in the presence of rat liver S9. Importantly, incubations with S9 in all cases decrease bioactivity in the EATS panel and pHBA was essentially devoid of significant biological activity.

Transcriptomics:

- All parabens elicited changes in the expression of a large number of genes in MFC7 cells, particularly at the highest dose tested.
- The parabens share a high degree of biological similarity across the category members, with a significant overlap in the affected pathways and the biological activity generally increasing with increasing chain length.
- A significant number of genes whose expression is up-or down-regulated by MP, EP or BP is also regulated in the same direction (up- or down-regulated) by PP (the target chemical).
- There are clearly fewer genes affected by pHBA; however, these were mostly different than the ones elicited by any of the parabens.
- These results demonstrate the overall biological similarity and thus the validity of the read-across category.
- The four parabens share an ability to up-regulate oestrogen response genes in MCF7 cells, indicating they share potential MOAs related to endocrine effects.
- The biological activity associated with the target chemical PP is most similar to that of BP. This correlates with the high structural similarity of these parabens (based on Tanimoto) and supports the read-across hypothesis that it is appropriate to use data from BP in read-across to PP.

Exposure Assessment:

- It was assumed that all cosmetic products contain all four parabens and that they are present at the maximum allowed concentration. The route of exposure is dermal.
- Aggregate exposure was calculated deterministically by multiplying the aggregate product exposure with the maximum concentration of the paraben allowed in each product according to regulation. The resulting estimates are very conservative.

- Aggregate exposure assessments were also calculated with the Creme Care model. This is a probabilistic exposure model and software for determining high-tier estimates of aggregate exposure to substances in personal care products and cosmetics. This is a refinement of the assessment of consumer exposure which is based on real consumer habits and practices.
- The external exposure according to four application scenarios using the refined assessment ranged between 0.059-0.368 mg/kg/day for MP; 0.019-0.262 mg/kg/day for EP; 0.014-0.154 mg/kg/day for PP; and 0.018-0.091 mg/kg/day for BP.

PBBK modelling:

- A PBBK model was developed and used to estimate the internal plasma concentrations of MP, PP and BP following whole body exposure in lotion
- The PoD dose of 2 mg/kg/day BP was administered by SC injection to rats. The C_{max} was 2.1 µ M, the AUC was 3.0 µ mole*h/L and the C_{avg} was 0.13 µ M.
- The internal exposure estimates for PP were C_{max} of 0.020 µ M, the AUC was 0.340 µ mole*h/L and the C_{avg} was 0.0140 µ M from the SCCS deterministic consumer exposure estimates;
- Th estimate of PP C_{max} was 0.0064 µ M, the AUC was 0.110 mmole*h/L and the C_{avg} was 0.0046 µ M from the Crème probabilistic (worst case) consumer exposure estimates
- The estimate of PP C_{max} was 0.00058 µ M, the AUC was 0.010 µ mole*h/L and the C_{avg} was 0.00042 µ M from the Crème probabilistic (realistic) consumer exposure estimates.

MoIE calculation:

- The MoIEs were calculated to be 284 using the SCCS deterministic values, and 887 and 9786 for Tier 2a and 2d consumer exposure estimates, respectively, using the Crème model probabilistic calculation.
- The MoIE was considered sufficient to support the use of PP in dermally applied cosmetic products.

5. Uncertainty assessment

In many of the methods and data presented herein, there are assumptions that are either inherent to the methodology or as a result of the decisions made during execution of the case study. For each of the types of data presented, an attempt has been made by the investigators to address uncertainty in their data reports. However, it can be useful to collectively consider the major assumptions and areas of resulting uncertainty in the risk assessment in order to determine the level of confidence and appropriate decision contexts in which it can be useful. Major assumptions and their impact on the assessment are listed below.

Table 9. Uncertainties in the NAMs used in the case studies

Data type/ Endpoint	How used [a]	Assumptions	Direction and Magnitude of Uncertainty [b]	Comments
In vivo data	WOE RAX RA	Good quality *in vivo* data performed according to international guidelines. Adverse effect has a dose response. PoD can be derived from the study.	++	*In vivo* study was chosen because was used by SCCS (2013). Study ranked Klimish score 3 (non-guideline, no dose-response).The PoD (NOEL) derived is 2 mg/kg/day for BP which is very conservative compared to other *in vivo* studies.
Exposure data	RA	Conservative (rounding, etc.) regarding aggregation	++	Reduces accuracy of values but increases confidence of the RA
NAM				
Molecular Docking/ER activity	WOE	These docking simulations can characterise the binding propensities of short linear chain parabens and their common ester hydrolysis metabolite pHBA towards twelve nuclear receptors	+/-	Docking simulations indicate a homogenous profile of weak activity with for the receptors considered by the Endocrine Disruptome tool, further substantiating the suitability of the four parabens to form one category
ToxCast/ Potency	WOE RA	ToxCast can increase confidence in the similarities of the structurally related chemicals in the category and inform on MoA & potency	+/-	MP and EP appear to have lower bioactivity in ToxCast assays than PP and BP, and pHBA did not demonstrate any significant activity in the assays. Based on ToxCast oestrogen receptor activity assays relative potency scaling factors could be derived. Uncertainty remains regarding the coverage of ToxCast assays, the metabolic capacity and the fact that no data on pHBA could be included in the analysis.
ADME Properties/pHBA activity	WOE RAX	pHBA, the main metabolite of parabens, does not contribute to the observed low reproductive toxicity potential associated with exposure to parabens	+/-	*In silico* predictions, EATS analysis and ToxCast evaluations differentiate pHBA from the parabens and support our assumption. pHBA toxicogenomics data demonstrated significantly less gene expression change as compared to the parabens (especially BP and PP). On the other hand, pHBA is not covered in the PBPK modelling and there is no estimate of internal exposure to pHBA which leaves some uncertainty.
CALUX assays/ER activity	WOE	Assay provides good quality data for the target and source chemicals on the oestrogen receptor binding and activation. The assay provides a potency trend among target and source compounds and positive control.	+/-	The assay was performed according to OECD TG by an experienced lab with track record of high reproducibility, low variability. CALUX assays are based using U2-OS cells, which have no endogenous receptors. This makes the assay highly specific and reduces the uncertainty. U2-OS cells have limited metabolic capacity, which might lead to false negative results if active metabolite would be produced *in vivo*, or false positive results if an active parent molecule would be readily metabolised *in vivo*. This uncertainty was reduced by performing the assays +/- liver S9 extract. Good quality data, with low potential to cause overestimation or underestimation
Toxicogenomics	WOE	Toxicogenomic data can inform on the gene changes and support the identification of the specific biologic activity of parabens.	+/-	The toxicogenomics studies were conducted under standardised conditions for the gene sets measured and for the cell type utilised with validated commercial transcriptional profiling platforms and statistical data analysis packages. While similar gene changes are observed in the MCF7 cells treated with parabens, but not pHBA, how

				these changes relate to *in vivo* effects is not known at this point. There is also uncertainty in the toxicogenomics data with regard to biological coverage because only one cell line was used.
PBBK	RA	PBBK model will provide the data on internal exposure of the target chemical based on different external exposure scenarios. Model will be used to calculate the internal exposure resulting from the PoD of the *in vivo* study.	+/-	A PBBK model was developed and used to estimate the internal plasma concentrations of MP, PP and BP following whole body exposure based on different exposure scenarios. The model has been previously published and validated. Internal exposure from the *in vivo* study was calculated. The ability to rely on a measure of internal rather than external exposure reduces the uncertainty in the risk assessment by incorporating chemical-specific information on the ADME parameters of the chemical in the experimental animal and the human. The rat SC injection dosing route has high uncertainty in the PBBK model because there are no rat SC kinetic data to address this uncertainty.

a How data was used in the case: RAX=read-across; RA=risk assessment; WOE=weight of evidence for biological similarity
b Key to direction and magnitude:
+, ++ = uncertainty results minor or major conservatism in the safety assessment (i.e. overestimation of risk).
-, - - = uncertainty results in minor or major concerns in the safety assessment (i.e. underestimation of risk).

6. Case study conclusion

6.1. Conclusion on PP itself based on this case study

The data provided herein constitutes the read-across justification for using *in vivo* reproductive toxicity data on the source chemical BP in read-across to fill the endpoint data gap for the target chemical PP. In addition, the other possible source chemicals MP and EP were included in the category approach to the read-across to evaluate similarity in bioactivity and explore potency trends across the category. This case study provides an example of integrating multiple data streams in an IATA approach to build a weight of evidence to support the appropriateness of reading across a PoD that can be used in confidence in risk assessment. While the *in vivo* reproductive toxicity data gap for PP in this case study was theoretical, the information gathered in this case study supports the safety of short linear chain parabens as used in cosmetic products, even when conservative assumptions are made in the safety assessment.

The parabens are generally substances of low toxicity and many of them reveal a NOAEL of up to 1000 mg/kg/day after repeated oral dosing. This is supported by new *in vivo* data on PP, generated to comply with EU REACh regulations, after daily oral administration of doses up to 1000 mg/kg to juvenile rats from the neonatal period (PND 4) through early adult life (PND 90) including uterotrophic assays and a full TK profile (Sivaraman *et al.*, 2018). There was no evidence of oestrogenic activity at any *in vivo* dose, and no effects on reproductive organs or function. The NOAEL was 1000 mg/kg/day. The predominant metabolite pHBA contributed to 95% of the total exposure at 1000 mg/kg/day. These data confirm the working hypothesis of this case study that all parabens are readily cleaved by esterase's and converted to the predominant metabolite, pHBA. A NOAEL of 1000 mg/kg/day as the highest dose tested was also identified in a 90-day repeated dose oral toxicity study in rats according OECD 408 and in a developmental toxicity study in rats according to OECD 414 (ECHA, 2018). Overall, there was no evidence of any adverse effects up to the limit dose of 1000 mg/kg/day.

Based on conflicting results from the literature, there are concerns that the parabens possess oestrogenic activity *in vivo*. However, the oestrogenic activity observed *in vitro* is extremely weak (approximately 3 to 6 magnitudes lower at maximum concentrations compared to 17beta-estradiol), and sporadic reports of *in vivo* oestrogenic effects of parabens appear to be very weak compared to dietary components or 17beta-estradiol. Therefore, although the parabens exhibit weak endocrine activity in *in vitro* test systems, the toxicological relevance for human safety is questionable. To date there is no *in vivo* evidence of adverse effects resulting from the weak endocrine activity of parabens. Furthermore, the safety assessment conducted in this case study for demonstration purposes resulted in margins of exposure for the parabens that would be considered protective for human health.

6.2. Conclusion on the usefulness of NAMs to support read across

As demonstrated in this case study, NAM data can provide useful information to facilitate the selection of the most appropriate analogue from a category of chemicals to read across to the target category member. In addition, NAMs can be used to investigate and inform on both the TK and TD properties of target and source chemicals in a given read-across

scenario and effectively establish their biological as well as the structural similarity. To summarise, ToxCast data showed the trend of oestrogen receptor activity for the target and source chemicals. Based on ToxCast data, the relative potency scaling factors were calculated. The calculated relative potency scaling factors were supported by other data from NAM such as *in silico* alerts, CALUX and Transcriptomics. These data also suggested the toxicological similarity between the target and source chemicals. The *in vitro* ADME data supported the toxicodyamic similarity between the target and source chemicals. PBPK modelling proved useful to estimate the internal plasma concentrations following the defined exposure scenario. Finally, MoIE was derived based on the estimated relative potency scaling factors and estimate internal concentrations of parabens in plasma. In conclusion, this case study provides an example of how NAM data can significantly add to the weight of evidence to support key aspects of a read-across hypothesis.

Appendix 1. *In vitro* metabolism (liver and skin and skin penetration of parabens)

Skin penetration using non-viable human skin

Method: The penetration of finite doses (10 µl/cm^2) of MP (3.21 µg/cm^2) and PP (2.51 µg/cm^2) into and through non-metabolically active (i.e. previously frozen) human skin was measured according to the OECD 426 and 28 Test Guidelines (OECD, 2004a, 2004b) using a standardised protocol (Jacques-Jamin *et al.*, 2017).

Results:

Figure 3. Profiles showing the rapid penetration of (A) MP and (B) PP into the receptor fluid after topical application to human skin.

Values are means of cumulative amounts in the receptor fluid from 4 donors (3 discs each). Following topical application in phosphate buffered saline (PBS), the initial penetration of both parabens into the receptor fluid was rapid and the amounts in the receptor fluid reached a plateau within 8 h. The penetration profile was similar for both parabens.

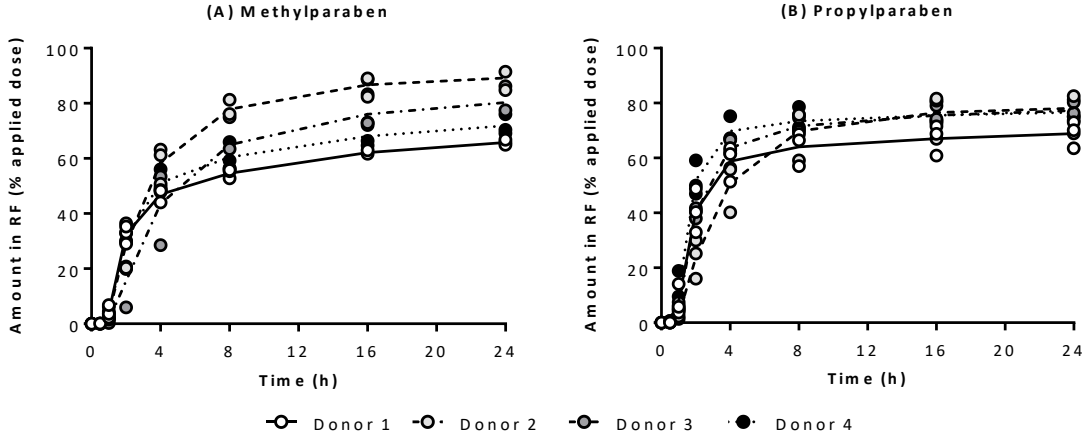


Table 10. Non-viable human skin penetration studies.

Dosing conditions and mass balance (A) and cutaneous distribution expressed as %applied dose (B) and µ g/cm^2 (C). With regard to vehicle effect, it is noted that the dermal delivery was decreased markedly when PP was applied in 100% ethanol compared to that when it was applied in PBS.

(A) Doses, solvent and mass balance

Chemical	Solvent	Dose (µg/cm^2)	Mass Balance
MP	0.01M PBS	3.21	94 ± 1.8
PP	0.01M PBS (Experiment 1)	2.42	95 ± 1.4
	0.01M PBS (Experiment 2)	2.51	97 ± 7.0
	100% ethanol (Experiment 2)	2.96	85 ± 5.7

(B) Cutaneous distribution - % Applied dose

Chemical	Dermal delivery	Skin wash	Stratum corneum	Epidermis	Dermis	Receptor fluid
MP	77 ± 9.7	17 ± 8.9	0.29 ± 0.13	0.14 ± 0.08	0.23 ± 0.23	77 ± 9.7
PP in PBS (Experiment 1)	76 ± 5.9	18 ± 6.3	1.51 ± 0.29	0.21 ± 0.10	0.23 ± 0.10	75 ± 5.9
PP in PBS (Experiment 2)	67 ± 10	30 ± 5.7	1.13 ± 0.7	0.40 ± 0.23	0.45 ± 0.24	65 ± 10
PP in ethanol (Experiment 2)	11 ± 6.0	73 ± 7.3	0.69 ± 0.60	0.32 ± 0.21	0.44 ± 0.25	10 ± 5.6

(C) Cutaneous distribution - µ g/cm^2

Chemical	Dermal delivery	Skin wash	Stratum corneum	Epidermis	Dermis	Receptor fluid
MP	2.47 ± 0.30	0.54 ± 0.29	0.009 ± 0.004	0.05 ± 0.003	0.007 ± 0.007	2.46 ± 0.3
PP in PBS (Experiment 1)	1.83 ± 0.15	0.43 ± 0.15	0.004 ± 0.007	0.005 ± 0.002	0.006 ± 0.002	1.82 ± 0.15
PP in PBS (Experiment 2)	1.66 ± 0.23	0.75 ± 0.15	0.03 ± 0.02	0.01 ± 0.01	0.01 ± 0.01	1.64 ± 0.23
PP in ethanol (Experiment 2)	0.33 ± 0.18	2.16 ± 0.19	0.02 ± 0.02	0.01 ± 0.01	0.01 ± 0.01	0.31 ± 0.17

Absorption and metabolism of PP in viable human skin explants

Method: The absorption and metabolism of PP was measured in viable human skin explants, as described by Genies *et al.* (2019). Briefly, human skin punches were seeded dermal side down in polycarbonate MD6 Inserts. The inserts were placed in a 6-well plate containing 1.5 mL culture medium and the skin explants incubated in an incubator at 37°C. The skin explants were incubated for 1 h before application of a finite dose of ^{14}C-PP (10 µ l/cm^2). Five doses (0.3 - 2.9 µ g/cm^2) and 4 kinetic times: 1, 2, 18 and 24 h were measured. The amount of ^{14}C-labelled chemical in the skin wash, skin and medium below the skin was quantified using liquid scintillation counting. Analysis of the chemicals and their metabolites in the medium was measured by radio-HPLC.

Results:

Table 11. Mass balance and % metabolism of PP in fresh viable human skin explants.

Data shows the extensive metabolism of PP over 24 h. Values are a mean (for mass balance ± SEM) from single skin disc from 4 donors.

	Dose applied (µg/cm^2)				
	0.3	0.5	1.0	2.4	2.9
Mass balance (%)	100 ± 11	92 ± 7	101 ± 10	97 ± 4	91 ± 4
Metabolites	99.51	99.59	99.70	99.90	99.86
Parent compound	0.49	0.42	0.30	0.11	0.14

Table 12. Amount of PP and its metabolites in the medium 24 h after topical application to human skin explants.

Amounts of metabolites in the medium expressed as a % of the applied dose (mean ± SEM) from single skin disc from 4 donors.

Analyte	% Applied dose
Parent compound + metabolites	56.0 ± 2.6
PP (parent compound)	0.2 ± 0.2
Total metabolites	55.8 ± 2.9
I – IX minor metabolites	2.4
X (4-Hydroxybenzoic acid (pHBA))	42.09 ± 2.8
XI (Not identified)	5.31 ± 0.9
XII (PP sulphate)	5.67 ± 0.3

Paraben biotransformation in primary human hepatocytes (PHH), human liver S9 and EpiSkin S9

S9 Method: Parabens were incubated in human liver S9 or reconstructed human epidermal model, EpiSkin™, both at 2 mg/ml, for up to 240 min (Eilstein *et al*., 2020). The main incubations contained a comprehensive mix of cofactors that allowed for some phase 1 reactions (892 µM NADPH) and several phase 2 reactions (520 µM Acetyl-CoA, 520 µM GSH, 372 µM UDP-glucuronic acid and 50 µM PAPS). Appropriate control incubations were included: (1) with S9 but no cofactors; (2) with cofactors but no S9 and (3) glass tubes.

PHH Method: The four parabens were incubated with primary human hepatocyte (PHH) suspensions and the depletion of parent compound, as well as the production of the main metabolite pHBA was quantified. All metabolites of each paraben were identified using LC-MS/MS

Results:

Figure 4. Depletion of parent paraben in the presence (closed symbols) and absence (open circles, dotted lines) of EpiSkin S9.

Values are a mean ± SD, expressed as a percentage of the initial concentration of paraben.

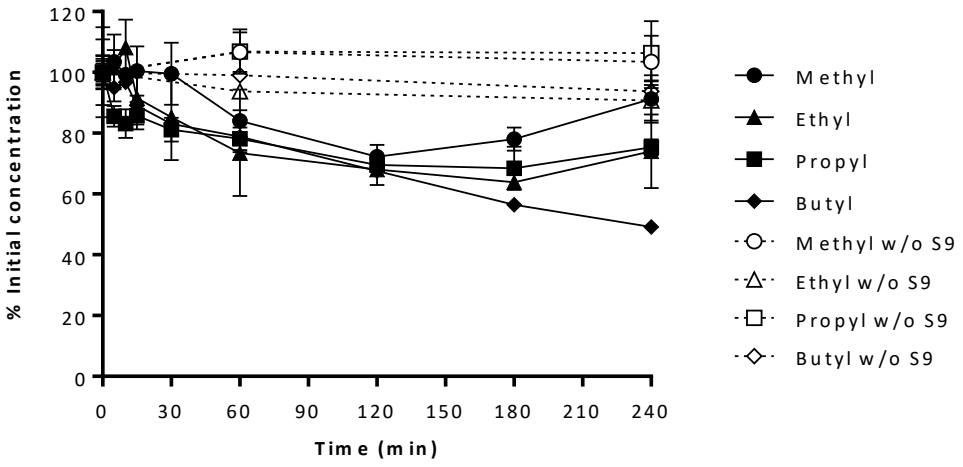


Table 13. $CL_{int,\ in\ vitro}$ values for parabens and pHBA incubated with primary human hepatocytes (PHH), human liver S9 and EpiSkin S9.

Incubations with PHH were run alongside reference compounds for high (naloxone), medium (midazolam) and low (tolbutamide) clearance compounds. Assuming 1 million PHH is equivalent to 1 mg S9, the $CL_{int,\ in\ vitro}$ values were comparable in PHH and liver S9. $CL_{int,\ in\ vitro}$ values were over 70-fold lower in EpiSkin S9 than liver S9.

Compound	PHH µL/min/million cells	Liver S9 µL/min/mg protein	EpiSkin S9 µL/min/mg protein
MP	94.0	129.3	0.68
EP	75.1	94.5	0.77
PP	91.3	105.1	0.5
BP	59.9	89.6	1.28
pHBA	0.41	Not metabolised	Not metabolised
Naloxone (high)	12.6	Not done	Not done
Midazolam (medium)	5.4	Not done	Not done
Tolbutamide (low)	1.6	Not done	Not done

Figure 5. Metabolism of parabens (A) and formation of pHBA (B) in primary human hepatocyte (PHH) suspensions.

The concentrations of parent paraben and 4-hydroxybenzoic acid (4-OH-benzoic acid) were measured using authentic reference standards. Values are a mean ± SD.

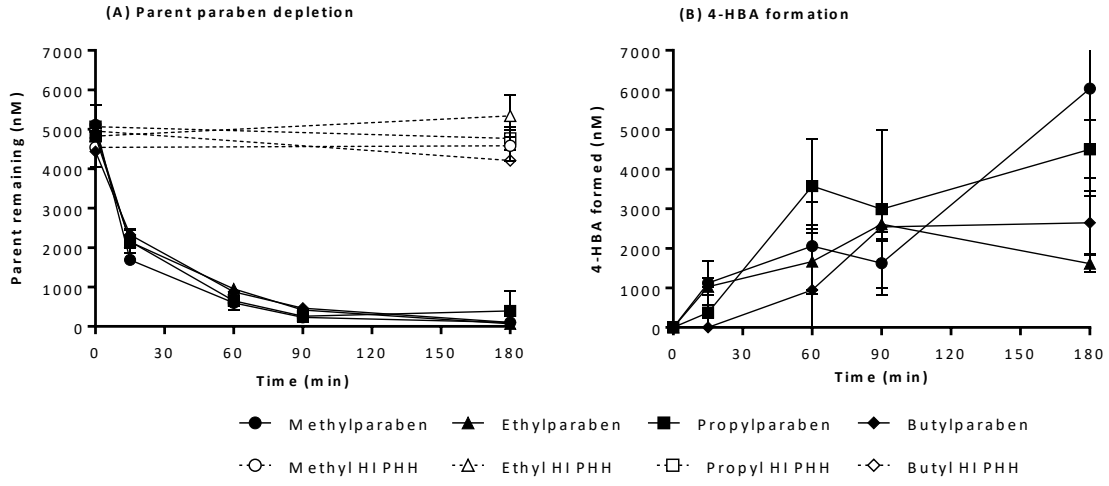


Table 14. Metabolites detected in incubations of 4 parabens and pHBA with primary human hepatocyte (PHH) suspensions, human liver S9 and EpiSkin S9.

Metabolites were detected using LC-MS/MS.

Compound	pHBA	Direct sulphate conjugate	Oxidised sulphate conjugate	Direct glucuronide conjugate	Oxidised glucuronide conjugate
MP					
Liver S9 / PHH	Present	Present	Present	Not detected	Not detected
EpiSkin S9	Present	Present	Not detected	Not detected	Not detected
EP					
Liver S9 / PHH	Present	Present	Present	Present	Not detected
EpiSkin S9	Present	Present	Not detected	Not detected	Not detected
PP					
Liver S9 / PHH	Present	Present	Present	Present	Not detected
EpiSkin S9	Present	Present	Not detected	Present	Not detected
BP					
Liver S9 / PHH	Present	Present	Present	Present	Present
EpiSkin S9	Present	Present	Not detected	Not detected	Not detected
pHBA					
PHH	NA	Not detected	Not detected	Not detected	Not detected
EpiSkin		Not tested	Not tested	Not tested	Not tested

Plasma protein binding

Method: Plasma protein binding studies were conducted using a cross-filtration method (modified from (Taylor & Harker, 2006)) using reference human plasma (fresh-frozen, pooled, mixed gender) spiked with test chemical. In addition, plasma samples with test chemical were incubated for 60 min to determine stability of the test chemical.

Results:

Table 15. Plasma protein binding (PPB) and stability of parabens (10 µM) in human plasma.

The % recovery of the parent chemical in the assay is also shown. Mean ± SD, n=3.

Paraben	PPB [%]	% Recovery	Stability control	
			% parent remaining after 1 h	nM 4 pHBA formed at 1 h (% parent metabolised)
MP	73.62 ± 1.92	92.0	96.4	158 ± 35 (1.6%)
EP	83.35 ± 0.54	90.9	94.3	109 ± 30 (1.1%)
PP	91.74 ± 0.06	88.0	85.6	550 ± 55 (5.5%)
BP	96.29 ± 0.25	82.7	97.3	0.0 ± 0.0 (0%)
pHBA	37.61 ± 3.13	101.1	97.9	NA
Warfarin (positive control)	97.86 ± 0.24	93.5	Not determined	NA

Appendix 2. Assays measuring interactions with oestrogen-, androgen-, thyroid signalling and steroidogenesis (EATS)

The use of *in vitro* assays to evaluate EATS focus on the following established methods that are validated and described in key publications, OECD test guidelines (TG) and detailed SOPs:

- E: (anti)Estrogens: ERα CALUX® (ER CALUX (OECD, 2016; van der Burg *et al.*, 2010a); EURL-ECVAM DB-ALM255 (JRC))
- A: (anti)Androgens: AR CALUX (van der Burg *et al.*, 2010b); ECVAM validation and OECD TG 258 incorporation/adaptation (Zuang *et al.*, 2018) and the EURL-ECVAM DB-ALM187 (JRC)
- T: Thyroid interference: TRβ CALUX, TTR* and TPO assay (Collet *et al.*, 2019; Zuang *et al.*, 2018); assay validation and OECD TG currently in preparation
- S: H295R steroidogenesis, coupled to CALUX read-out according to OECD TG 456 (OECD, 2011)

Oestrogens and androgens

CALUX bioassays are derived from human osteosarcoma cells (U2-OS), that are modified to respond specifically to a pathway of choice. Cells that are exposed to compounds of interest not only express proteins that are under normal circumstances associated to these pathways, but also luciferase. By addition of the appropriate substrate for luciferase, light is emitted. The amount of light produced is proportional to the amount of ligand-specific pathway activation (or pathway inactivation, in the case of an antagonistic response), which is benchmarked against a relevant reference compound. To be able to correct for possible non-specific effects, including cytotoxicity of the samples under investigation, the cytotox CALUX bioassay is used. The cytotox CALUX bioassay constitutively expresses luciferase and hence, light is constantly emitted. A dose-depend reduction of emitted light is indicative for cytotoxic effects of the compound under investigation. In some cases, activation occurs which can be related to luciferase stabilisation (Sotoca *et al.*, 2010).

Thyroid interferences

In addition to establishing direct interferences with receptor activation established with the TRβ CALUX, interferences via transporter proteins and hormone synthesis pathways play a role in EDC disruption of the thyroid axis. Disruption of the plasma transport of the thyroid hormone T4, an endpoint that is particularly relevant in foetal- and brain tissue, is determined using a TTR-binding assay with the TRβ CALUX bioassay as read-out. The TTR-binding assay is based on binding competition between a fixed concentration of T4 and dilution series of test compounds. Increasing concentrations of test items capable of competing with T4 for TTR-binding sites will result in a decreased amount of TTR-bound T4. Following separation of TTR-bound and free T4, the amount of TTR-bound T4 is determined using the TRβ CALUX bioassay.

Thyroid hormone synthesis in the thyroid gland is catalysed by thyroid peroxidase (TPO). TPO binds iodine to the tyrosine residues which ultimately become the thyroid hormone molecules. Compounds that inhibit TPO activity prevent iodothyronine production in the thyroid gland. The luminol-based hTPO inhibition assay is based on the oxidation of

luminol by hydrogen peroxide during which light is emitted. This reaction is catalysed by hTPO. In the presence of compounds inhibiting the hTPO catalytic activity, the amount of light emitted is reduced.

Steroidogenesis

The *in vitro* H295R Steroidogenesis Assay (H295R) with oestrogenic and androgenic read-out and described in OECD test guideline 456, utilises a human adenocarcinoma cell line (NCI-H295R cells). The assay has been developed and standardised as a screen for chemical effects on steroidogenesis, specifically the production of 17beta-estradiol and testosterone. The goal of the assay according to the OECD guideline is to provide a YES/NO answer with regard to the potential of a chemical to induce or inhibit the production of 17beta-estradiol and testosterone; however, it is also possible to report quantitative results by determining a lowest observed effect concentration (LOEL). The induction or inhibition of the synthesis of 17beta-estradiol and testosterone is determined using the AR CALUX and ERα CALUX bioassays respectively.

Methods:

Preparation of stock solutions and determination of solubility

All parabens were dissolved in DMSO and diluted in DMSO. The highest stock concentration was dictated by the compound's solubility in assay medium. Solubility of the dilution series was evaluated by preparing standard dilutions in CALUX assay medium (0.1% DMSO), TPO incubation buffer (1% DMSO), and TTR incubation buffer (3.2% DMSO). For the standard CALUX assays (anti-) ERα, (anti-) AR, (anti-)TRβ and Cytotox, a dilution series of 13 concentrations (in 0.5 Log M increments) was exposed to the cells at 0.1% DMSO; for H295R, TPO and TTR incubations, five concentrations (in 1.0 Log M increments) were selected based on compound toxicity and solubility.

Determination of cytotoxicity

Prior to evaluation of the reaction mixture on the various bioassays, the compounds were tested for cytotoxicity using the U2OS based CALUX cytotox bioassay, by exposing cytotox CALUX cells to serial dilutions of the compounds in DMSO. The cytotox CALUX cells constitutively express luciferase. Exposure of the cytotox CALUX cells to compound causing cytotoxicity will result in a reduction of luminescence. Compound concentrations causing 20% reduction of luminescence are considered cytotoxic.

hTPO inhibition assay

hTPO was derived from Nthy-ori 3-1 cells. Cell lysate containing hTPO in Glycine-NaOH buffer (pH 9.0) was incubated for 30 minutes at 37 ºC in the presence of serial dilutions of the compounds in DMSO (1% compound stock in incubation mixture). The incubation mixture was transferred to 96-well microtiter plates after which luminol (34.8 μM) and H_2O_2 (1.7 mM) were added. Luminescence was measured on a Berthold luminometer.

TTR-binding assay

Serial dilutions of the compounds were incubated in Tris-buffer (pH 8.0) overnight at 4 ºC in the presence of TTR (0.058 μM) and a fixed concentration of T4 (0.052 μM) (3.2% compound stock in incubation mixture). After incubation, TTR-bound and free T4 were separated on a Bio-Gel P-6DG column. The eluate was added to assay medium after which TRβ CALUX cells were exposed for 24 hours (see *CALUX bioassays*).

H295R steroidogenesis assay

To determine the effect of the compounds on sex hormone synthesis, H295R cells were seeded in 24 wells plates in assay medium and exposed for 48h to a serial dilution of the compounds in triplicate. To quantify the levels of 17beta-estradiol and testosterone produced by the H295R cells after exposure, the assay medium was analysed on the ERα and AR CALUX bioassays as described below (*CALUX bioassays*). In brief, the H295R cells are exposed to serial dilutions of the compounds for 48 hours, and subsequently the levels of testosterone and 17β-estradiol produced by the cells are determined. Changes in hormone levels compared to a vehicle control exposure indicates that certain enzymes involved in steroidogenesis were being affected by the test compound.

CALUX bioassays

For determination of the (anti-)ERα, (anti-)AR and (anti-)TRβ CALUX activities, CALUX cells were seeded in assay medium. Following exposure of the CALUX cells to serial dilutions of the compounds in triplicate for 24h, the induction of luciferase production is quantified by measuring luminescence (Berthold luminometer) following addition of the substrate luciferin. On each plate, a complete calibration curve for each respective bioassay is also analysed using the relevant reference compounds. For the analyses in the presence of rat liver S9, 0.03% (vv) of PB/BNF induced rat liver S9 mix (including NADPH and an NADPH-regenerating system) was added to the cells directly after addition of the compounds, which remained there for the duration of the exposure period (24h). For more details on the metabolic activation, see reference (van Vugt-Lussenburg *et al.*, 2018).

Data *analysis*

Analysis results of the parabens, expressed as induction relative to the standard reference compound, are interpolated in the calibration curves of each respective bioassay for quantitative determination of disruptive potential using the statistical software package GraphPad Prism V5.03. Only dilutions that do not show any signs of cytotoxicity (relative induction in the cytotox CALUX bioassay > 80%) are used for final evaluation of analysis results. All analysis results are expressed in molar (M) final in well.

Results:

Figure 6. Cytotox CALUX results.

Luciferase activity (% of vehicle control) is plotted against compound concentration (LogM) final in well. The assay was performed in the absence (blue) and presence of metabolic enzymes (rat liver S9 fraction). The threshold of activity (20% decrease of activity, PC20) is indicated as a dotted line.

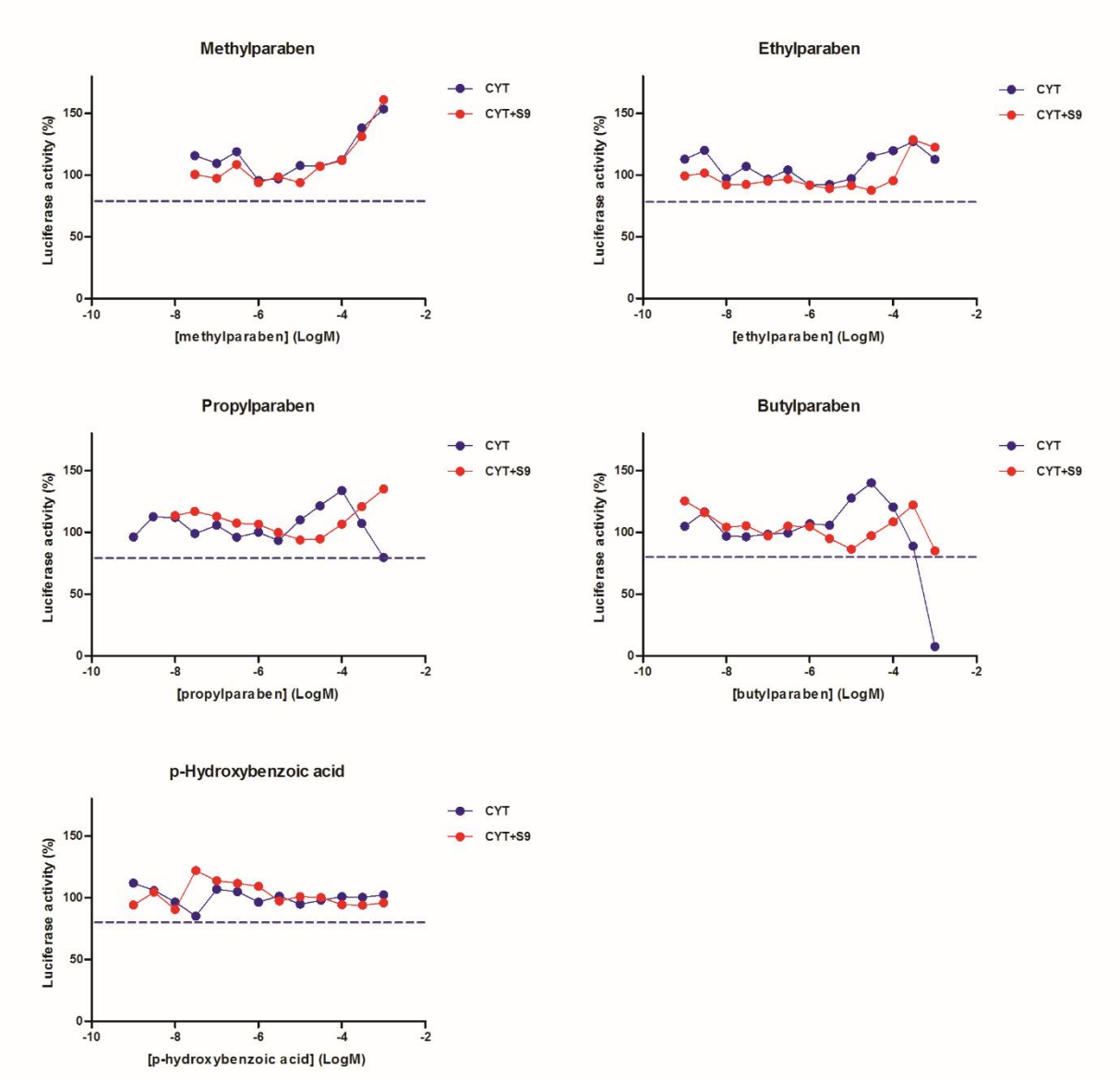


Figure 7. ERα CALUX results.

Receptor activation (% of maximum) is plotted against compound concentration (LogM) final in well. The assay was performed in the absence (blue) and presence of metabolic enzymes (rat liver S9 fraction). The threshold of activity (10% activity compared to reference compound 17β-estradiol (E2), PC10) is indicated as a dotted line. The reference curve is presented in black (no S9) and purple (with S9).

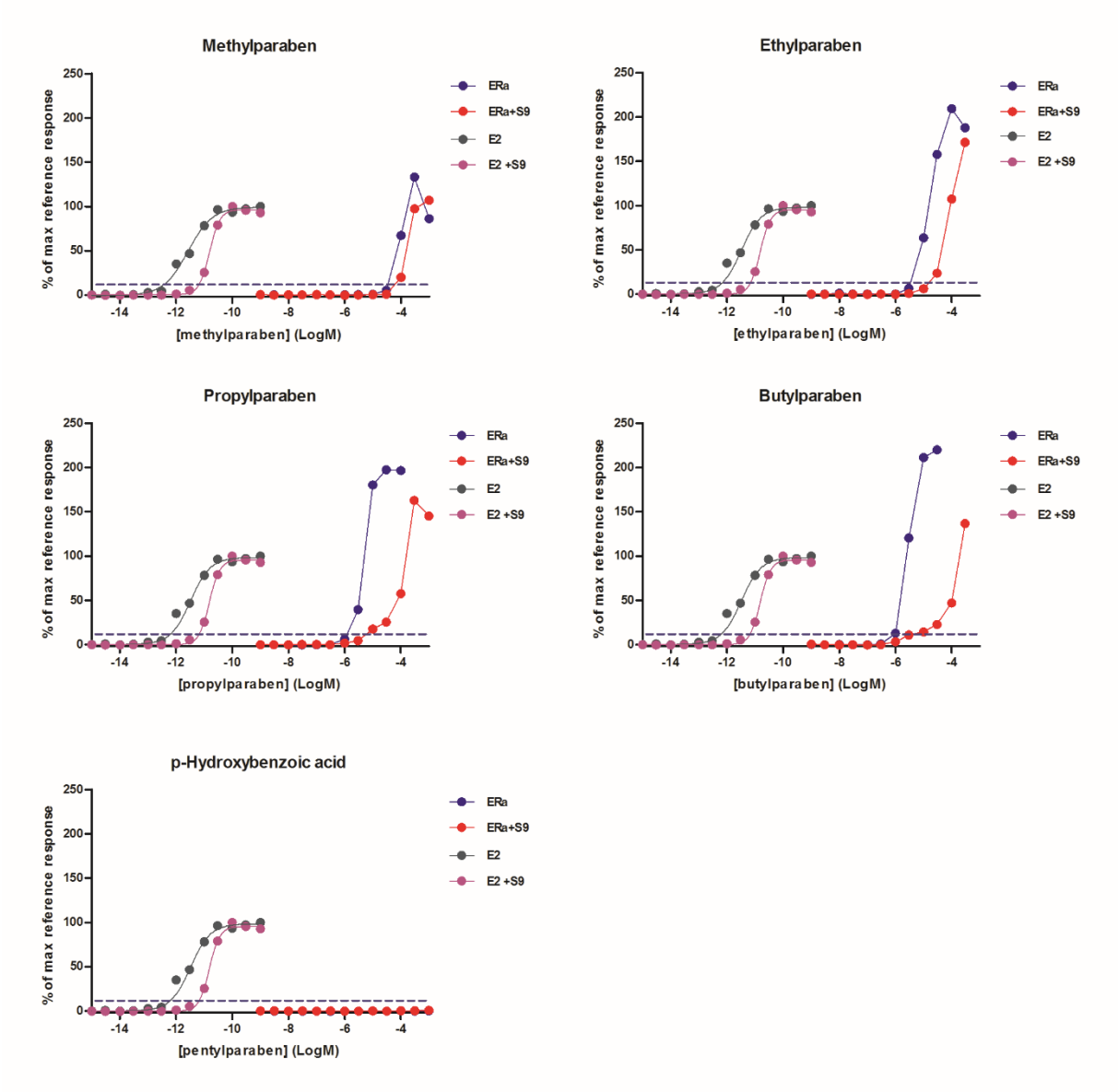


Figure 8. Anti-AR CALUX results.

Activity (% compared to EC50-agonist response) is plotted against compound concentration (LogM) final in well. The assay was performed in the absence (blue) and presence of metabolic enzymes (rat liver S9 fraction). The threshold of activity (20% inhibition of activity, PC20) is indicated as a dotted line. The reference curve is presented in black (no S9) and purple (with S9).

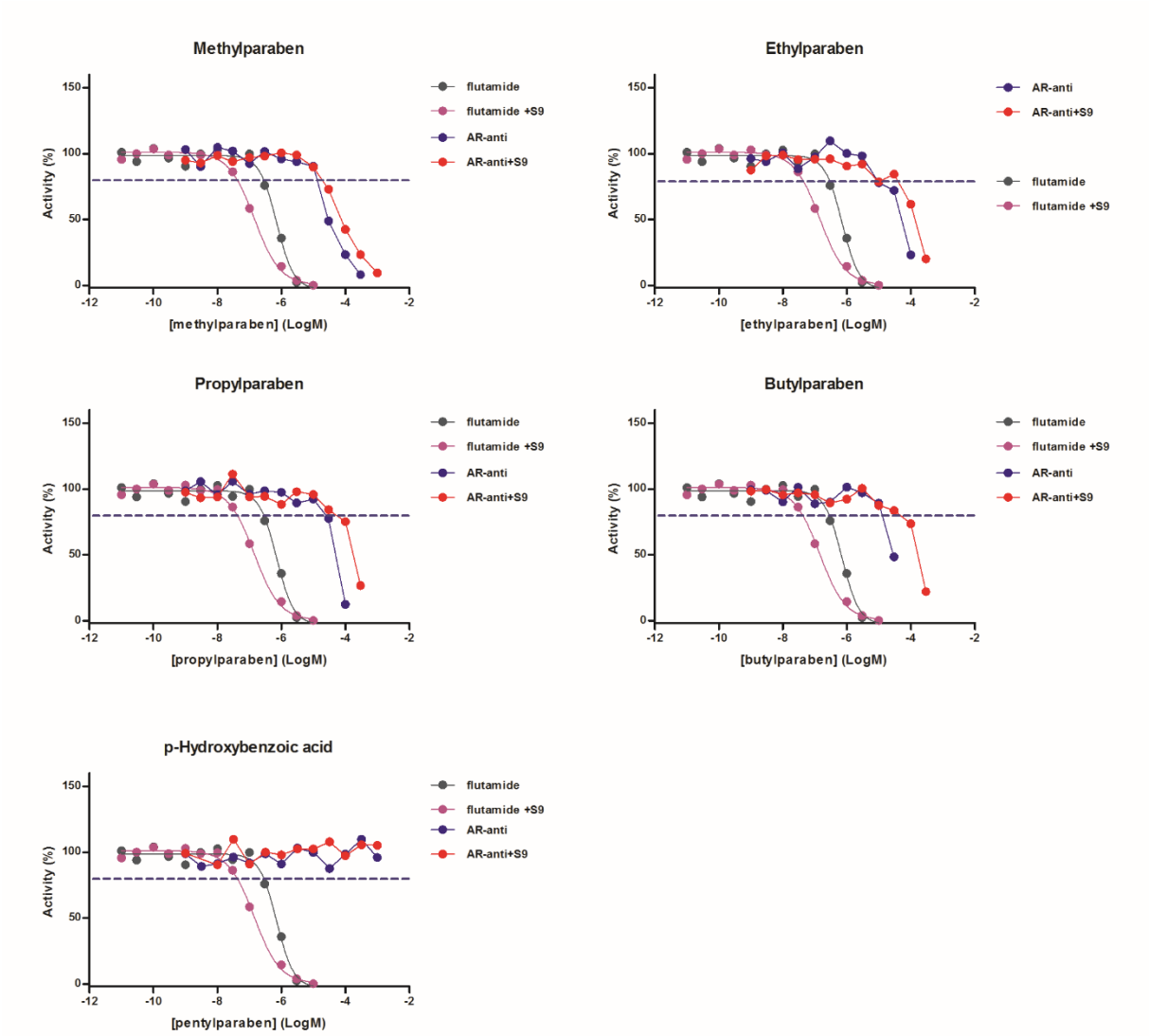


Table 16. Summary of EATS testing results.

PC10 (for agonistic tests)/PC20 (for antagonistic tests) values are shown in -Log M; the colour indicates the potency (yellow < orange < red).

End point	MP		EP		PP		BP		pHBA	
	-S9	+S9	-S9	+S9	-S9	+S9	-S9	+S9	-S9	+S9
Cytotoxicity	>	>	>	>	-3.5	>	-4.0	-3.0	>	>
(Anti-) oestrogenic and (anti-) androgenic assays										
ERα CALUX	-4.5	-4.2	-5.5	-4.8	-6.0	-5.1	-6.0	-5.0	>	>
anti-ERα CALUX	>	>	>	>	>	>	>	>	>	>
AR CALUX	>	>	>	>	>	>	>	>	>	>
anti-AR CALUX	-4.9	-4.7	-4.7	-4.4	-4.5	-4.2	-4.9	-4.3	>	>
Thyroidogenic assays										
TRβ CALUX	>	>	>	>	>	>	>	>	>	>
anti-TRβ CALUX	-3.0	>	>	>	>	>	>	>	>	>
TTR	-2.7	nd	-4.8	nd	-4.5	nd	-4.8	nd	>	nd
hTPO	-2.0	nd	>	nd	>	nd	>	nd	-3.0	nd
Steroidogenesis										
H295R-E2	-5.0	nd	-5.0	nd	-5.0	nd	-5.0	nd	-3.0	nd
H295R-T	>	nd	-4.0	nd	-4.0	nd	>	nd	>	nd

Appendix 3. Transcriptomics to evaluate biological similarity of the parabens

Method: The MCF7 cells were exposed for 6 hours to vehicle control or to three non-cytotoxic concentrations of MP, EP, PP and BP or pHBA (1, 50 and 500 µ M). Three independent cell cultures per dose of each paraben or its metabolite were used as biological replicates. Total RNA was extracted from each individual cell culture biological replicate after 6 hours of exposure, and samples from each replicates of each treatment group with high quality mRNA were processed for transcriptional profiling. The transcriptional response of the MCF7 cells to vehicle control, or to the appropriate dose of each of the parabens tested, was evaluated using a comprehensive transcriptional profiling platform, TempO-Seq (BioSpyder). Differentially expressed genes were identified for each of the parabens at each dose of exposure using the statistical analysis of the log2 transformed data using Limma software, which uses Benjamin and Hochberg's method to control the false discovery rate (FDR) at 5% across all samples including controls. Differentially expressed genes were those that met the threshold of change defined as an FDR > 0.05 and a fold change of ≥ 1.2.

For the pathway enrichment analyses, the Molecular Signatures Database (MSigDB v7.0) was utilised, since it contains 22596 gene sets grouped within 50 Hallmark gene sets. The Hallmark gene sets summarise and represent specific well-defined biological states or processes and display coherent expression. These gene sets were generated by a computational methodology based on identifying overlaps between gene sets in other MSigDB collections and retaining genes that display coordinate expression (http://software.broadinstitute.org/gsea/msigdb/collection_details.jsp; accessed on September 10, 2019).

Results:

Figure 9. Heat map of the genes whose expression was modified in MCF-7 cells after exposure to parabens or pHBA.

Up-regulated genes in red; down-regulated genes in blue.

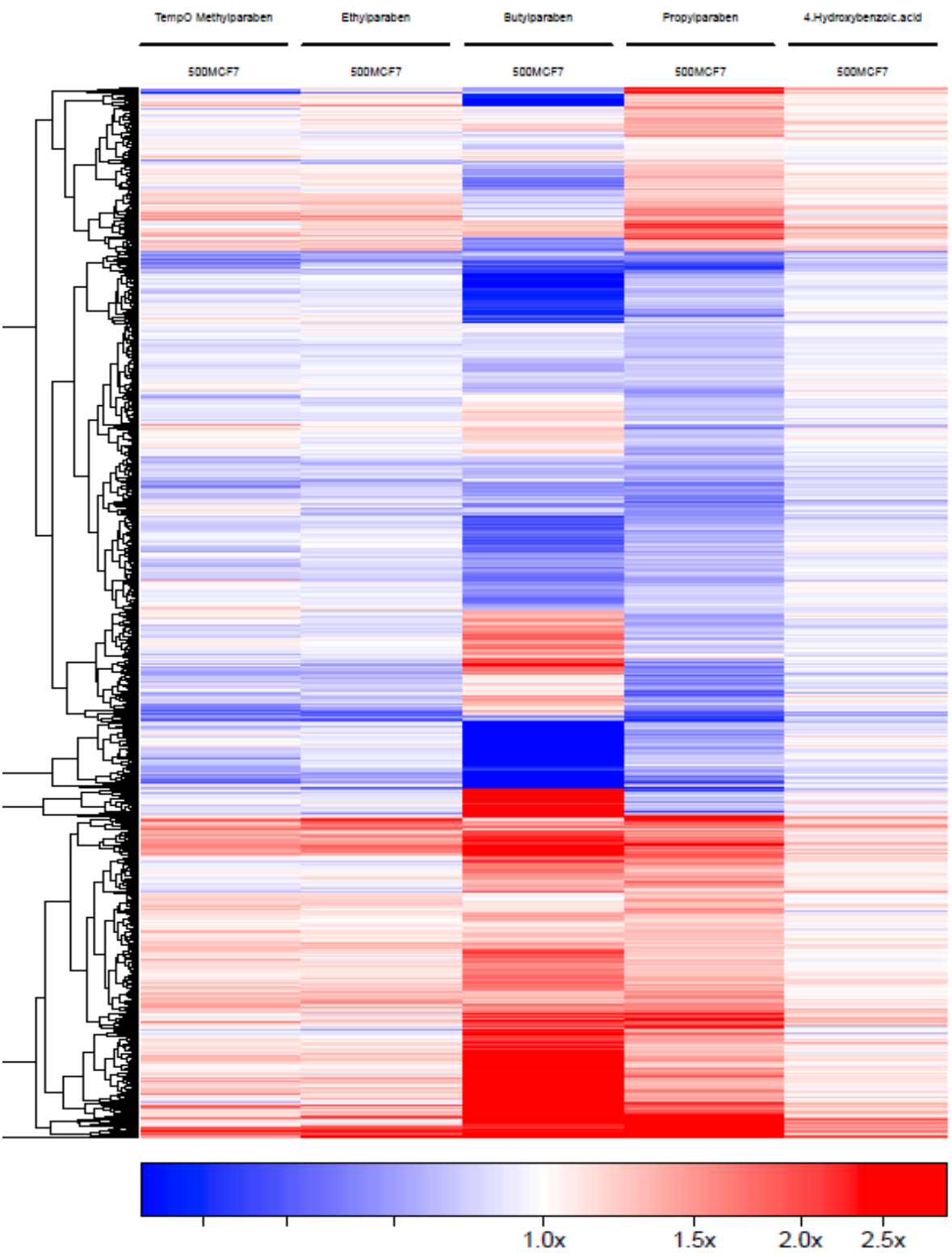


Table 17. Principal pathways affected by MP.

Identified using the Molecular Signatures Database (MSigDB v7.0).

BioSpyder Gene Set Name [# Genes (K)]	Description	# Genes in Overlap (k)	p-value	FDR q-value
HALLMARK_ESTROGEN_RESPONSE_EARLY [200]	Genes defining early response to oestrogen.	29	5.76 e^{-39}	2.88 e^{-37}
HALLMARK_ESTROGEN_RESPONSE_LATE [200]	Genes defining late response to oestrogen.	24	2.86 e^{-30}	7.15 e^{-29}
HALLMARK_UNFOLDED_PROTEIN_RESPONSE [113]	Genes up-regulated during unfolded protein response, a cellular stress response related to the endoplasmic reticulum.	7	1.24 e^{-7}	2.07 e^{-6}
HALLMARK_UV_RESPONSE_UP [158]	Genes up-regulated in response to ultraviolet radiation.	6	1.72 e^{-5}	2.15 e^{-4}
HALLMARK_GLYCOLYSIS [200]	Genes encoding proteins involved in glycolysis and gluconeogenesis.	6	6.43 e^{-5}	4.59 e^{-4}
HALLMARK_IL2_STAT5_SIGNALING [200]	Genes up-regulated by STAT5 in response to IL2 stimulation.	6	6.43 e^{-5}	4.59 e^{-4}
HALLMARK_P53_PATHWAY [200]	Genes involved in p53 pathways and networks.	6	6.43 e^{-5}	4.59 e^{-4}

Table 18. Principal pathways affected EP.

Identified using the Molecular Signatures Database (MSigDB v7.0).

BioSpyder Gene Set Name [# Genes (K)]	Description	# Genes in Overlap (k)	p-value	FDR q-value
HALLMARK_ESTROGEN_RESPONSE_EARLY [200]	Genes defining early response to oestrogen.	34	1.21 e^{-44}	6.03 e^{-43}
HALLMARK_ESTROGEN_RESPONSE_LATE [200]	Genes defining late response to oestrogen.	29	5.18 e^{-36}	1.29 e^{-34}
HALLMARK_UNFOLDED_PROTEIN_RESPONSE [113]	Genes up-regulated during unfolded protein response, a cellular stress response related to the endoplasmic reticulum.	16	3.86 e^{-20}	6.43 e^{-19}
HALLMARK_MTORC1_SIGNALING [200]	Genes up-regulated through activation of mTORC1 complex.	11	1.05 e^{-9}	1.05 e^{-8}
HALLMARK_TNFA_SIGNALING_VIA_NFKB [200]	Genes regulated by NF-kB in response to TNF [GeneID=7124].	11	1.05 e^{-9}	1.05 e^{-8}
HALLMARK_GLYCOLYSIS [200]	Genes encoding proteins involved in glycolysis and gluconeogenesis.	9	1.99 e^{-7}	1.66 e^{-6}
HALLMARK_EPITHELIAL_MESENCHYMAL_TRANSITION [200]	Genes defining epithelial-mesenchymal transition, as in wound healing, fibrosis and metastasis.	8	2.32 e^{-6}	1.16 e^{-5}

Table 19. Principal pathways affected by PP.

Identified using the Molecular Signatures Database (MSigDB v7.0).

BioSypder Gene Set Name [# Genes (K)]	Description	# Genes in Overlap (k)	p-value	FDR q-value
HALLMARK_ESTROGEN_RESPONSE_EARLY [200]	Genes defining early response to oestrogen.	53	$2.2\ e^{-42}$	$1.1\ e^{-40}$
HALLMARK_TNFA_SIGNALING_VIA_NFKB [200]	Genes regulated by NF-kB in response to TNF [GeneID=7124].	48	$2.77\ e^{-36}$	$6.92\ e^{-35}$
HALLMARK_UNFOLDED_PROTEIN_RESPONSE [113]	Genes up-regulated during unfolded protein response, a cellular stress response related to the endoplasmic reticulum.	34	$1.05\ e^{-29}$	$1.75\ e^{-28}$
HALLMARK_ESTROGEN_RESPONSE_LATE [200]	Genes defining late response to oestrogen.	36	$6.1\ e^{-23}$	$7.62\ e^{-22}$
HALLMARK_P53_PATHWAY [200]	Genes involved in p53 pathways and networks.	35	$6.4\ e^{-22}$	$6.4\ e^{-21}$
HALLMARK_HYPOXIA [200]	Genes up-regulated in response to low oxygen levels (hypoxia).	31	$5.43\ e^{-18}$	$4.52\ e^{-17}$
HALLMARK_MTORC1_SIGNALING [200]	Genes up-regulated through activation of mTORC1 complex.	28	$3.21\ e^{-15}$	$2.29\ e^{-14}$

Table 20. Principal pathways affected by BP.

Identified using the Molecular Signatures Database (MSigDB v7.0).

BioSpyder Gene Set Name [# Genes (K)]	Description	# Genes in Overlap (k)	p-value	FDR q-value
HALLMARK_ESTROGEN_RESPONSE_EARLY [200]	Genes defining early response to oestrogen.	46	$7.96\ e^{-61}$	$3.98\ e^{-59}$
HALLMARK_ESTROGEN_RESPONSE_LATE [200]	Genes defining late response to oestrogen.	36	$1.89\ e^{-43}$	$4.73\ e^{-42}$
HALLMARK_UNFOLDED_PROTEIN_RESPONSE [113]	Genes up-regulated during unfolded protein response, a cellular stress response related to the endoplasmic reticulum.	16	$3.91\ e^{-18}$	$6.51\ e^{-17}$
HALLMARK_MYC_TARGETS_V1 [200]	A subgroup of genes regulated by MYC - version 1 (v1).	13	$1.35\ e^{-10}$	$1.69\ e^{-9}$
HALLMARK_MTORC1_SIGNALING [200]	Genes up-regulated through activation of mTORC1 complex.	12	$1.75\ e^{-9}$	$1.75\ e^{-8}$
HALLMARK_HYPOXIA [200]	Genes up-regulated in response to low oxygen levels (hypoxia).	10	$2.22\ e^{-7}$	$1.85\ e^{-6}$
HALLMARK_EPITHELIAL_MESENCHYMAL_TRANSITION	Genes defining epithelial-mesenchymal transition, as in wound healing, fibrosis and metastasis.	9	$2.16\ e^{-6}$	$1.08\ e^{-5}$

Appendix 4. Exposure assessment for dermally applied cosmetics

For the purpose of this case study it was assumed that all cosmetic products contain all four parabens and that they are present at the maximum allowed concentration. The route of exposure is dermal. Based on their chemical structure, parabens belong to Cramer Class I for which the TTC value is 0.03 mg/kg/day.

Low tier (conservative) and high tier (more realistic) estimates of aggregate external exposure for parabens in cosmetics were conducted. The low tier (Tier 1) assessment of consumer exposure was conducted using habits and practices data for the products to be considered for aggregate exposure, according to the Scientific Committee on Consumer Safety (SCCS) Notes of Guidance for Cosmetic Testing (SCCS, 2018). Then, in order to refine the Tier 1 estimate, a number of different data sources and the Creme Care aggregate model were utilised in a systematic manner to perform a high tier (Tier 2), refined estimate, of consumer exposure.

Tier 1

The Tier 1 exposure assessment assumes a maximum EU regulatory use level in all products:

- For the max concentration of MP and EP, regulation states the maximum is 0.4% (as acid), which equates to 0.44% as ester for MP and 0.48% as ester for EP
- For the max concentration of PP and BP, regulation states the maximum is 0.14% (as acid), which equates to 0.18% as ester for PP and 0.19% as ester for BP.

The products considered are shown in **Table 21**, with their external product exposures.

Table 21. Product exposures for the deterministic calculation of aggregate exposure for preservatives through cosmetic use (SCCS, 2018).

Type of cosmetic product exposure	*Product category*	*Exposure product ($E_{product}$) (g/d)*	*$E_{product}$ normalized by body weight[1] (mg/kg bw/d)*
Rinse-off skin& hair cleansing products	Shower gel	0.19	2.79
	Hand wash soap	0.20	3.33
	Shampoo	0.11	1.51
	Hair conditioner	0.04	0.67
Leave on skin& hair cleansing products	Body lotion	7.82	123.20
	Face cream	1.54	24.14
	Hand cream	2.16	32.70
	Deodorant non-spray	1.50	22.08
	Hair styling	0.40	5.74
Make-up products	Liquid foundation	0.51	7.90
	Make-up remover	0.50	8.33
	Lipstick	0.06	0.90
	Eye make-up	0.02	0.33
	Mascara	0.025	0.42
	Eyeliner	0.005	0.08
Oral care Products[2]	Toothpaste	0.14	2.16
	Mouthwash	2.16	32.54
TOTAL		17.4	269

[1]The specific body weight of the persons involved in the study is used and not the default value of 60kg
[2]Oral care product categories are not corrected and are presumed here to only represent dermal exposure (mucosa)

Aggregate exposure was calculated deterministically by multiplying the aggregate product exposure (17.4 g/day or 269 mg/kg) with the maximum concentration of the paraben allowed in each product according to regulation. The resulting estimates (see Table 22) are very conservative.

Table 22. Tier 1 deterministic aggregate exposure estimates to parabens through cosmetic use (SCCS 2013).

Paraben	Maximum Concentration Allowed (% as ester)	External Exposure (mg/day)	External Exposure (mg/kg/day)
MP	0.44%	76.56	1.18
EP	0.48%	83.52	1.29
PP	0.18%	31.32	0.48
BP	0.19%	33.06	0.51

Tier 2

A number of refinements are possible when performing a higher-tier exposure assessment, by examining the various factors driving the conservative nature of a screening level type exposure assessment. The refinement introduced at this point was the use of a probabilistic model (Crème Care) with which to assess consumer exposure that is based on real consumer habits and practices, as opposed to deterministically summing the contribution

from each product category. Other assumptions regarding product retention remain consistent as in the first case Tier 1 assessment.

Moving to a probabilistic and subject-oriented model can provide refinement of the estimates of exposure (although not always), but also offers a framework with which to introduce further inputs that can be used to improve estimates of exposure. This probabilistic modelling methodology allows:

- The use of statistical distributions to characterise substance concentrations
- The use of occurrence to account for the presence of chemicals in some, but not all products.

Crème Care

Aggregate exposure assessments were calculated with the Creme Care model. This is a probabilistic exposure model and software for determining high-tier estimates of aggregate exposure to substances in personal care products and cosmetics. It is built upon a habits and practices database of over 36,000 consumers from a product use survey developed by Kantar Worldpanel for Europe and the United States (US), detailing frequency of product use, co-use and site of application for cosmetic categories over a seven-day period. Amount per application data is based upon clinical studies for the same products, which are in the form of statistical distributions. Additional required parameters for exposure estimates such as bodyweight, height and skin surface areas are also included in the form of statistical distributions from published sources (e.g. NHANES) and standard calculations (e.g. the Dubois formula). The model calculates aggregate exposure to a chemical via the dermal, inhalation and oral routes, with systemic exposure expressed on an absolute or per unit bodyweight basis and dermal exposure as per unit of skin surface area by site of application.

The model works by combining data on the concentration of a substance within each product category with the data in the habits and practices database. Concentration values can be point estimates or statistical distributions, described empirically or parametrically, and with or without occurrence (i.e. the likely occurrence in each product category). Daily exposures are simulated for each individual consumer based on selected inputs, which are used to calculate distributions of chronic or acute exposure in the population being assessed, which can be stratified by age, gender, or geography. The calculated exposure distribution (described using the appropriate measure to compare with the reference dose in question) is in turn described using appropriate statistics and can be broken down to assess the relative contribution of each product category to exposure, or alternatively to assess the relative magnitude of the exposure at each application site.

Four scenarios were considered:

2a. Paraben always present, max concentration in regulation

2b. Paraben always present, concentration at current use range according to Cosmetic Europe Product Preservation Survey (2016)

2c. Using paraben occurrence data according to Mintel GNPD, max concentration in regulation

2d. Using paraben occurrence data according to Mintel GNPD, concentration at current use range according to Cosmetic Europe Product Preservation Survey (2016).

Of these, scenario 2d represents a realistic exposure scenario for paraben exposure through use of cosmetic products.

The assessments were done in an EU population consisting of 26,209 individuals. In order to have a like-with-like comparison with the SCCS notes of guidance (SCCS, 2018) and to examine the impact of various refinements, retention was kept consistent with tier 1. The product categories were also the same in both tiers, which are based upon those that were in the original Kantar Worldpanel survey.

Paraben Concentration

The Cosmetics Europe Expert Team on Product Preservation launched a comprehensive survey across its membership on preservative use in cosmetic products in 33 European countries in early 2017, which gathered data on paraben concentration and occurrence (according to volume) in cosmetic products. The survey had a high level of response and participation from not only Cosmetics Europe membership, but from the wider EU Cosmetics Industry. This dataset holds data on over 47000 formulations.

Paraben Occurrence

Occurrence was derived using to methods. Firstly, from the Products Preservation survey of 2016, which recorded data on over 470,000 formulation, the volume of formulas containing parabens versus those that did not was used to infer occurrence.

Secondly, occurrence data were derived from an online available database called Mintel GNPD from 2008 through 2016, which lists the labelling information for a large number of products in a number of cosmetic categories.

Mintel GNPD (http://portal.mintel.com/) is an online database that tracks consumer product launches across the globe and provides reliable data on ingredient labelling. The database is divided into categories such as Face/Neck Care, Body Care, Eye Care, and Lip Care. Using the site's search function returns the total number of products in that category. Adding the ingredient name (paraben) gives the subset of those products containing this ingredient.

To derive the occurrence of parabens in the products used in this assessment, the total number of products in each category per year was counted, and the proportion of these that contain the paraben based on whether it was listed on the label or not was used to calculate the occurrence. Note that this approach therefore assumes equal market shares; if a product that does or does not contain the paraben has a large market share then the occurrence in reality will be reduced or increased accordingly. However, for simplicity an equal market share approach is used initially as market shares or sales volumes are not readily available.

To reduce uncertainty and err on the side of conservatism, occurrence was inferred from the two datasets (survey and Mintel data from 2008 through 2016). The highest occurrence value from the two data sets was taken and rounded as follows:

- if occurrence is < 1%, then round up to 2%
- if occurrence is < 2%, then round up to 5%
- if occurrence is < 5%, then round up to 10%
- if occurrence is < 10%, then round up to 20%
- if occurrence is > 10% but below 30% then round up to 50%
- if occurrence is > 30%, then assume 100%

Summary of Exposure Assessment Results

As a distribution of exposure is the resulting output of the assessment, characterising this distribution can be done in a number of ways. In the following, the 95th percentile of exposure is used to represent the upper exposure in the population. Additionally, when using real habits and practices data, exposure statistics can be calculated over two populations. The first is the Total Population, i.e. all subjects in the survey, and the other is the Exposed Population. The Exposed population is the more conservative, because it removes zero exposures (where individuals are using products that do not contain parabens), and so is used in this risk assessment.

The Tier 2 assessment P95 results for the four scenarios in the Exposed Population are provided in Table 23.

Table 23. Tier 2 probabilistic aggregate estimates of exposure to parabens through cosmetic use.

	External Exposure (mg/kg/day)			
Paraben	Tier 2a	Tier 2b	Tier 2c	Tier 2d
MP	0.368	0.111	0.183	0.059
EP	0.262	0.059	0.078	0.019
PP	0.154	0.053	0.07	0.014
BP	0.091	0.037	0.045	0.018

Uncertainty in Tier 2 Assessment

Several factors potentially influence the uncertainty of the aggregate exposure assessment. Uncertainties that can be identified around factors of the data are listed in Table 24 with the direction of the effect (overestimation/underestimation potential) noted.

Table 24. Qualitative evaluation of the influence of uncertainties on the estimate of exposure.

Assumption	Direction and Magnitude of Uncertainty	Comments
Frequency of use data in Creme Care from the Kantar diary (France, Germany, Spain and Great Britain) for cosmetic products of interest (excluding eye shadow, mascara, face powder, blusher, eyebrow pencil, make up remover, eye cream) representing frequency of use data for these products amongst all EU citizens.	+/-	Good quality data, with low potential to cause overestimation or underestimation
Frequency of use data in Creme Care from the Netherlands (Biesterbos, 2013) for eye shadow, mascara, eyebrow pencil, make-up remover representing frequency of use data for these products amongst EU.	+/-	Good quality data, with low potential to cause overestimation or underestimation
Frequency of use data in Creme Care from P&G for eye cream representing frequency of use data for these products amongst EU citizens.	+/-	Good quality data, with low potential to cause overestimation or underestimation
Amount per use data in Creme Care from the Cosmetics Europe studies for cosmetics and additional data from the CTFA and Biesterbos studies (2013) for certain cosmetics representing amount per use data for these products amongst EU.	++/- -	Sufficiently representative data but there may be some gaps in knowledge for specific countries which could mean these data overestimate or underestimate the average for the EU population
The Product Preservation Survey provides data on paraben concentration and occurrence that represents cosmetic products sold across the European Union.	+/-	Good quality data, with low potential to cause overestimation or underestimation
The Mintel GNPD database contains a comprehensive list of cosmetic products sold across the European Union.	+/-	Good quality data, with low potential to cause overestimation or underestimation
Occurrence from the GNPD database is representative of the likelihood of consumer use of cosmetic products.	+/-	Good quality data, with low potential to cause overestimation or underestimation
The potential for brand loyalty (Arcella *et al* 2003), e.g. to always buy the same product containing a high % of ingredient.	-	High end users and upper end concentrations have been included, but 100% high end use by an individual is not assumed.

* Key to direction and magnitude:
+, ++, +++ = uncertainty likely to cause small, medium or large over-estimation of exposure.
-, - -, - - - = uncertainty likely to cause small, medium or large under-estimation of exposure.

Appendix 5. Physiologically-based biokinetic modelling in exposure assessment

The physiological structure of PBBK models provides a particularly useful framework for conducting cross species extrapolations (Clewell & Andersen, 1985). A number of recent reviews of PBTK/PBPK modelling in environmental risk assessment are available (Campbell *et al.*, 2012; Clewell *et al.*, 2014; Clewell & Clewell, 2008). A typical equation for a perfusion-limited tissue describes the mass balance for the uptake and clearance of the chemical in the tissue, in this case the liver:

$$dA_{Liver}/dt = Q_L * (C_{Arterial} - C_{Venous}) - Cl_{Liver}$$

This equation can be interpreted as: The rate of change in the mass of the chemical (A_{Liver}) in the liver is equal to the liver blood flow (Q_L) multiplied by the difference between the concentrations in the blood entering and leaving the liver ($C_{Arterial} - C_{Venous}$), minus the metabolic clearance of the chemical in the liver (Cl_{Liver}). These models typically rely on three types of parameters; physiological (e.g., tissue volumes, blood flows), physicochemical (e.g., octanol:water portioning, vapour pressure, water solubility), and biochemical (e.g., absorption rates, metabolism, clearances). The particular parameters needed depend on factors such as the chemical properties and the purpose of the model. Various guidance documents for the application, use, and reporting of PBBK models have been published (U.S. EPA, 2006; U.S. FDA, 2018; WHO, 2010).

The necessary physiological parameters (tissue weights, blood flows, ventilation rate) for a number of mammals (mouse, rat, dog and human) are available in the literature (Brown *et al.*, 1997), and parameters for other species can be estimated allometrically (Lindstedt & Schaeffer, 2002). Tissue:blood partition coefficients for a chemical can be estimated using quantitative structure-property relationships (Peyret *et al.*, 2010), while the clearance of the chemical in different species can be determined by *in vitro* studies with hepatocytes or cellular fractions and incorporated into the PBBK model using *in vitro* to *in vivo* extrapolation (Yoon *et al.*, 2012).

A PBBK model for MP, PP, and BP was previously developed by Campbell *et al.* (2015). and applied in this case study. Available (in the literature) skin penetration parameters from Cross and Roberts (2000) were used rather than values from *in vitro* ADME data generated from NAMs presented in this case study. (In this case study, the *in vitro* ADME data is solely used to compare the behaviour of the four parabens in the category using similar *in vitro* conditions and to provide information helpful in the selection of the most appropriate source chemical for the read-across.)

Methods:

PBBK model:

The PBBK model for MP, PP and BP, developed by Campbell *et al.* (2015). The model has perfusion-limited compartments for skin, liver, gastrointestinal tract (GI), fat, blood, and a lumped compartment representing the rest of the body. Exposure is characterised in the skin (dermal), stomach (oral), and blood (intravenous, IV). Tissue blood flows and volumes are set to values of Brown *et al.* (1997). Tissue:blood partition coefficients (PC) were estimated by Campbell *et al* (2015). using the unified algorithm of Peyret *et al.* (2010). The

skin:blood PC was estimated as a weighted average of liver and fat PCs (i.e., 0.7*PL+0.3*PF).

Metabolism is described via first-order hydrolysis in liver, skin, and GI, and conjugation through glucuronidation or sulfation in liver. The hydrolysis product, pHBA, and total conjugated parabens were modelled using a single, whole body compartment, with elimination via urinary excretion.

Oral exposure is modelled using a 3-compartment absorption model, including stomach, duodenum, and remaining GI tract tissues. First-order absorption to the GI tissue occurs in the stomach and duodenum, and hydrolysis occurs in the GI tissue compartment. The blood flow from the GI enters the liver via the portal vein.

Dermal exposure is modelled using a skin surface compartment to house the applied dose, and a single skin compartment. Metabolism and transfer to the systemic circulation occur in the skin compartment. The skin is separated into exposed and "rest of skin" compartments. Dermal absorption is driven by a permeation rate from the surface into the skin (units of 1/cm^2 /hr), the fraction absorbed, the exposure area, and the amount of paraben applied. Transfer from the skin to the blood is modelled assuming a well-mixed, blood flow-limited exposed skin compartment and a skin:blood partition coefficient.

A schematic of the model structure is shown in Figure 10. A list of physiological parameter values is shown in Table 27and chemical-specific parameters for the modelled paraben compounds are shown in Table 25. A description of the model calibration conducted by Campbell *et al.* (2015) and any modifications made for the current modelling effort are described below.

Figure 10. PBBK model schematic for parabens.

Parent compound may be hydrolysed in the liver, skin, and gastrointestinal (GI) tissue, and conjugated (glucuronidation and sulfation) in the liver. Parent and metabolites may be excreted in urine. A fat compartment is included as a storage tissue.

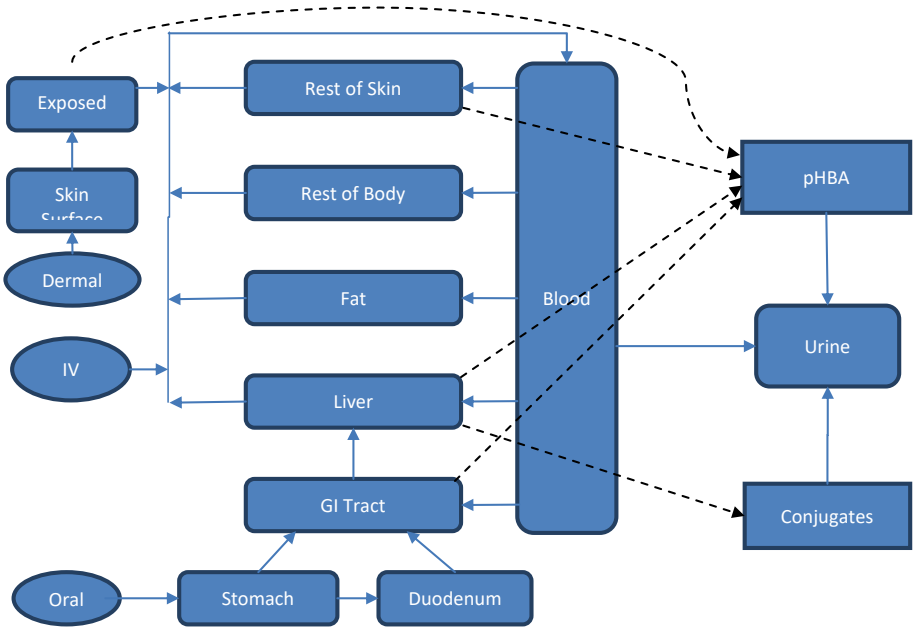


Table 25. Human PBBK physiological parameter values used for the paraben model.

The source of parameter values is Campbell *et al.* (2015). The GI cytosolic mg of protein per g tissue was used for the skin (no data available for skin). The rat skin microsomal mg of protein per g tissue was used for human skin (no human data available).

Parameter	Units	Symbol	Value
Body mass	kg	BW	60
Skin thickness	cm	Depth	0.1
Blood Flows (Fraction of Cardiac Output)			
Cardiac Output	$L/h/kg^{0.75}$	QCc	15
Fat	1	QFc	0.052
Gastrointestinal Tract	1	QGIc	0.181
Liver	1	QLc	0.046
Skin	1	QSkC	0.058
Volumes (fraction of BW)			
Blood	1	VBldC	0.079
Rest of Body	1	Vdc	0.92
Fat	1	VFc	0.21
Gastrointestinal Tract	1	VGIc	0.014
Liver	1	VLc	0.026
Urine Production	ml/kg/d	VolUc	22
Protein Content (per g tissue)			
Microsomal			
Liver	mg	MGMPGL	40
Gut	mg	MGMPGGI	23.6
Skin	mg	MGMPGSK	1.93
Cytosol			
Liver	mg	MGCPGL	76
Gut	mg	MGCPGGI	1.6
Skin	mg	MGCPGSK	1.6

Table 26. Chemical-specific human PBBK parameters.

The source of parameter values is Campbell *et al.* (2015).

Parameter	Units	Symbol	Value			
			MP	EP	PP	BP
Molecular weight	g/mol	MW	152.15	166.176	180.2	194.23
Partition Coefficients						
Fat		PF	1.0	2.2	5.1	9.2
GI		PGI	0.94	1.5	2.5	3.6
Liver		PL	0.94	1.5	2.5	3.6
Rest of body		PVd	0.70	0.99	1.6	2.3
Urinary clearance						
Paraben	$L/h/kg^{0.75}$	CLUC	0.13	0.090	0.034	0.0013
conjugate	$L/h/kg^{0.75}$	CLUGC	0.31	0.31	0.31	0.31
pHBA	$L/h/kg^{0.75}$	CLUPC	0.31	0.31	0.31	0.31
Dermal absorption						
Fraction available	1	FracAvail	0.16	0.16	0.16	0.16
Permeability	$1/h/cm^2$	P	3.2E-06	8.8E-06	1.0E-05	8.8E-06
Metabolism						
Microsomal Hydrolysis						
Liver	L/h/mg protein	KalMc	1.9E-01	1.5E-01	3.3E-02	1.8E-02
Skin	L/h/mg protein	KASKMc	2.4E-04	2.5E-04	6.0E-05	4.9E-05
GI	L/h/mg protein	KagiMc	4.2E-04	6.0E-04	2.6E-03	5.8E-03
Cytosolic Hydrolysis						
Liver	L/h/mg protein	KalCc	1.4E-02	8.7E-03	6.2E-06	2.6E-03
Skin	L/h/mg protein	KaSkCc	3.3E-04	3.2E-04	2.4E-06	9.2E-05
Conjugation						
Liver	L/h/mg protein	KalGMc	5.9E-01	1.6E-01	4.3E-02	2.4E-03

Parameter Estimation by Campbell et al. (2015)

Hydrolysis rates for human liver and skin tissue were similarly calculated using *in vitro* measurements in microsomal and cytosolic incubations (Jewell *et al.*, 2007a; 2007b). The rate of hydrolysis in the GI compartment was calculated using *in vitro* measurements in human small intestinal microsomes (Ozaki *et al.*, 2013). The reaction velocities (nmol/min/mg protein), protein content in incubation, and volume of incubation from these studies were used to calculate an intrinsic clearance (L/h/mg protein) for human liver, skin, and GI by fitting a prediction of the change in incubation concentration of paraben based on these parameters. Ozaki *et al.* (2013) did not provide cytosolic hydrolysis rates for the GI tissue, and this potential contribution to hydrolysis was assumed to be negligible for the dermal exposure scenario to be simulated.

The rates of direct conjugation and urinary excretion (along with dermal penetration and fraction absorbed) of BP in the liver were fitted to *in vivo* human data in Janjua *et al.* (2008), as shown in Figure 11. Janjua *et al.* exposed 25 human males to an approximate whole-body dermal application of cream containing 2% BP by weight. Cream was applied for 5 consecutive days, and 24-hour urine output was analysed for BP and total (free plus conjugated) paraben, controlling for a pre-exposure background level in urine. The conjugation and urinary excretion of MP and PP were also fitted to human data, using the ratio of free to total paraben levels in 100 urine samples (Ye *et al.*, 2006) and 16 serum samples (Ye *et al.*, 2009) from biomonitoring studies. Urinary excretion rates of pHBA and

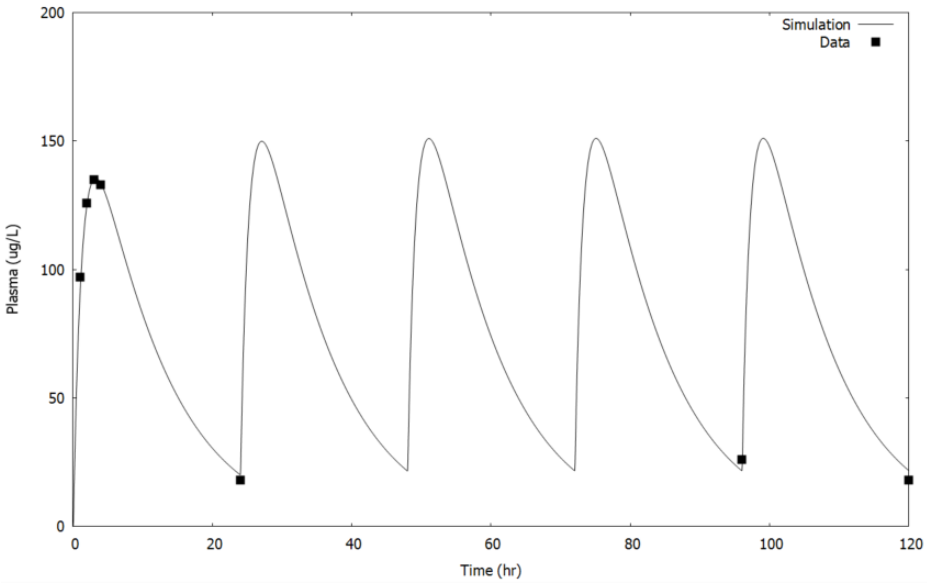
conjugates were set to the glomerular filtration rate for all parabens.

BP PBBK simulations compared to human data of Janjua *et al.* (2008). Subjects were exposed daily to an approximately whole-body application of ointment containing BP for 5 days. BP in plasma and urine were determined daily (Figure 11).

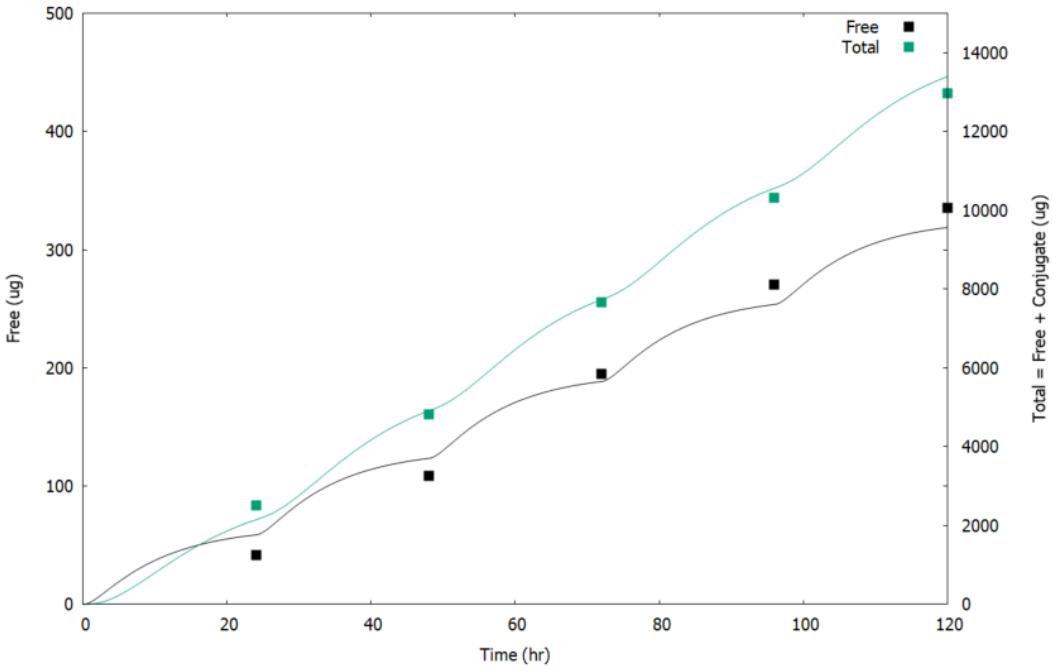
Figure 11. BP PBBK simulations compared to human data.

BP in (A) plasma and (B) urine were measured and predicted.

(A) Plasma



(B) Urine

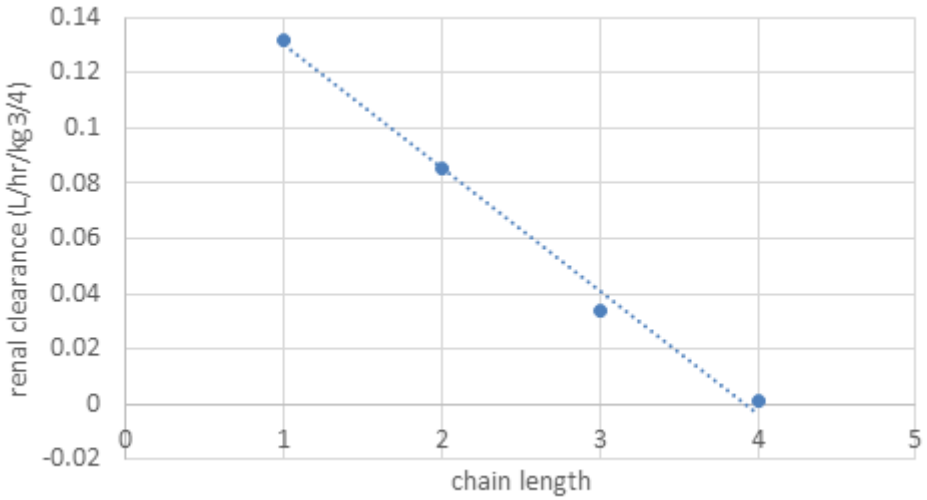


Dermal uptake of MP and PP was estimated using a parallelogram approach. The ratio of the fitted value of the dermal penetration rate, and the *in vitro* value of penetration rate measured by Cross & Roberts (2000) was applied to *in vitro* determined values of penetration rate for MP and PP. The fraction available for absorption fitted to BP (16%) data was used for all parabens.

Model Modifications

The ability to simulate EP was added, using the same model structure as for the other parabens. An approach similar to that of Campbell *et al.* (2015) for the other parabens was used to estimate the PBBK parameters for consistency. The hydrolysis in the liver, skin, and GI was calculated based on the data in Jewell *et al.* (2007a; 2007b) and Ozaki *et al.* (2013). The dermal penetration parameter was set to the value for PP, based on the measurement of similar penetration rates for EP and PP reported by Cross & Roberts (2000). There were no serum data for EP in the Ye *et al.* (2009) biomonitoring study, though there were urine free and bound determinations in Ye *et al.* (2006). The direct conjugation rate in the liver was adjusted to reproduce the 7% free/bound fraction, and the urinary clearance rate was estimated by interpolating between the values derived for the other parabens, which followed a linearly decreasing trend with increasing chain length (see Figure 12).

Figure 12. Correlation of chain length and renal clearance rate used to estimate EP urinary clearance.



To be consistent, the partition coefficients (PC) were recalculated for all parabens using GastroPlus. The inputs to the algorithm were the octanol:water partition coefficient (Kow), the fraction unbound in plasma (fup), and the dissociation constant (pKa). Kow and pKa were obtained from PubChem, and fub was according to the values in **Table 14**. This altered the previous simulations of MP, PP and BP very little, since the new PC estimates were very similar to those of Campbell *et al.* (2015). Hydrolysis rates for the human GI in Ozaki *et al.* (2013) were corrected. A miscalculation of protein content in the *in vitro* incubation were found upon examination of the files used by the original authors of the PBBK model. Also, the cytosolic hydrolysis rates were included in the recalculation of paraben kinetics. These changes amount to relatively small changes in the predicted kinetics, as the liver cytosolic rates of hydrolysis are low compared to microsomal, the skin

penetration is rapid and results in extensive first-pass skin metabolism, and the gastrointestinal metabolism has less effect on the simulated dermal exposures simulated here than would be potentially impactful on an oral exposure.

A route for SC injection dosing was added to the model to enable simulation of plasma exposure from a NOEL for BP SC toxicity in the rat reported by Fisher *et al.* (1999). The SC dose was modelled as a bolus input directly to the exposed skin compartment. Release from the exposed skin compartment is currently driven by skin to blood partitioning and blood flow to the skin, without consideration for release from the exposure vehicle. Skin metabolism was not applied to the SC dose compartment since the injection would have been administered below the metabolically active skin layers. Rat physiological and chemical-specific model parameters are shown in Table 27 and Table 28, respectively.

Table 27. Rat PBBK physiological parameter values used for the paraben model.

The source of parameter values is Campbell *et al.* (2015).

Parameter	Units	Symbol	Value
Body mass	kg	BW	0.23
Blood Flows (Fraction of Cardiac Output)			
Cardiac Output	$L/h/kg^{0.75}$	QCc	15
Fat		QFc	0.07
Gastrointestinal Tract		QGIc	0.151
Liver		QLc	0.024
Skin		QSkC	0.058
Volumes (fraction of BW)			
Blood		VBldC	0.074
Rest of Body		Vdc	0.92
Fat		VFc	0.07
Gastrointestinal Tract		VGIc	0.027
Liver		VLc	0.034
Urine Production	ml/kg/d	VolUc	57
Microsomal Protein Content (per g tissue)			
Liver	mg	MGMPGL	45
Gut	Mg	MGMPGGI	2.1

Table 28. Chemical-specific rat PBBK parameters.

The source of all parameter values except hydrolysis rates is Campbell *et al.* (2015). The hydrolysis rates were recalculated from the data of Ozaki *et al.* (2013).

Parameter	Units	Symbol	Value
Molecular weight	g/mol	MW	194.23
Partition Coefficients			
Fat		PF	7.4
GI		PGI	1.9
Liver		PL	1.9
Rest of body		PVd	1.8
Urinary clearance			
Paraben	$L/h/kg^{0.75}$	CLUC	0.001
conjugate	$L/h/kg^{0.75}$	CLUGC	0.24
pHBA	$L/h/kg^{0.75}$	CLUPC	0.24
Hydrolysis			
Liver	L/h/mg protein	KalMc	1.8E-02
Skin	L/h/mg protein	KASKMc	4.9E-05
GI	L/h/mg protein	KagiMc	5.8E-03
Conjugation			
Liver	L/h/mg protein	KalGMc	2.4E-03

Results:

Figure 13. Human PBBK simulation of internal plasma concentration of MP, EP, PP and BP following whole body exposure in lotion.

Body surface area = 13845 cm^2 (whole body, 60 kg human). MP exposure = 1.18 mg/kg/d, EP exposure = 1.29 mg/kg/d, PP exposure = 0.48 mg/kg/d and BP exposure = 0.51 mg/kg/d.

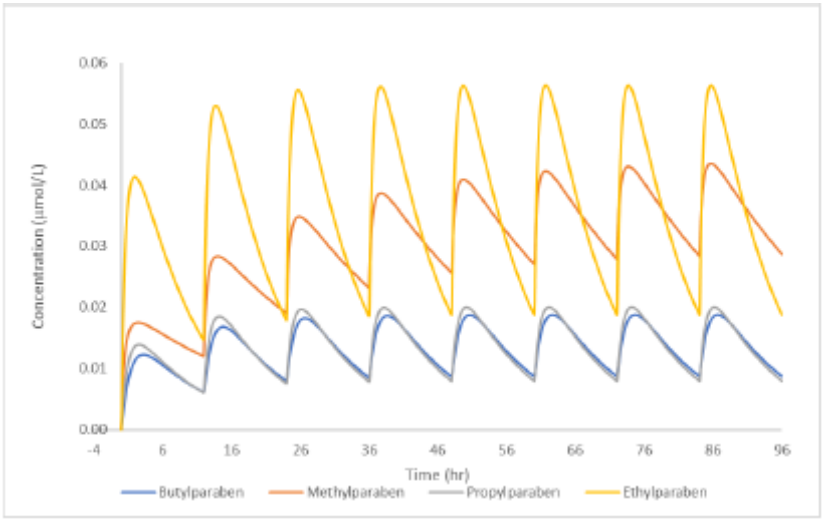


Table 29. Summary of human plasma data for the simulations shown in Figure 13.

Exposure estimates based on the Tier 1 deterministic assessment.

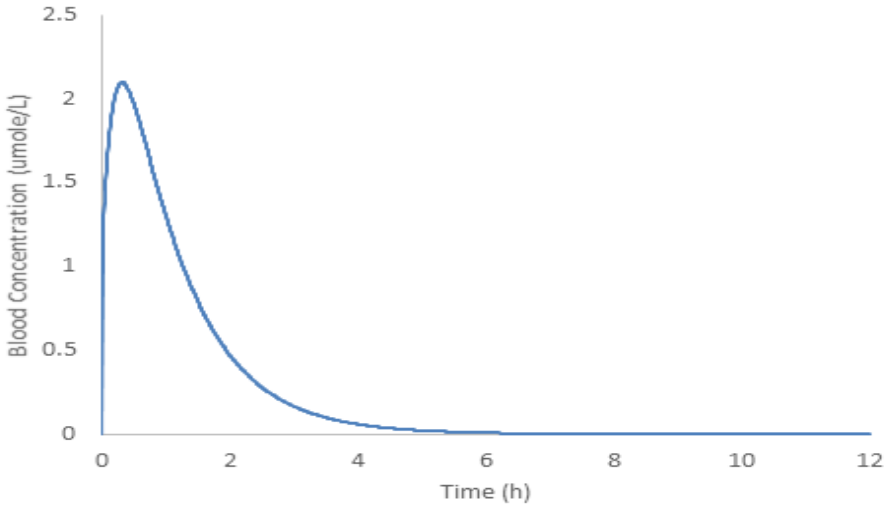
	Exposure mg/kg/d	C_{max} µmole/L	AUC µmole*hr/L	C_{avg} µmole/L
MP	1.18	4.4E-02	8.9E-01	3.7E-02
EP	1.29	5.6E-02	9.1E-01	3.8E-02
PP	0.48	2.0E-02	3.4E-01	1.4E-02
BP	0.51	1.9E-02	5.2E-04	2.2E-05

Table 30. Summary of human plasma data for the simulations of exposures estimated with the Crème Global exposure tool.

Chemical	Exposure (Crème Global)			C_{max}	AUC	C_{avg}
	Tier	mg/kg/d	µg/cm²	µmole/L	µmole*h/L	µmole/L
MP	2a	0.368	0.80	1.4E-02	2.8E-01	1.2E-02
EP	2a	0.262	0.57	1.1E-02	1.8E-01	7.7E-03
PP	2a	0.154	0.33	6.4E-03	1.1E-01	4.6E-03
BP	2a	0.091	0.20	3.4E-03	6.1E-02	2.5E-03
MP	2b	0.111	0.24	4.1E-03	8.4E-02	3.5E-03
EP	2b	0.059	0.13	2.6E-03	4.2E-02	1.7E-03
PP	2b	0.053	0.11	2.2E-03	3.8E-02	1.6E-03
BP	2b	0.037	0.08	1.4E-03	2.5E-02	1.0E-03
MP	2c	0.183	0.40	6.8E-03	1.4E-01	5.8E-03
EP	2c	0.078	0.17	3.4E-03	5.5E-02	2.3E-03
PP	2c	0.07	0.15	2.9E-03	5.0E-02	2.1E-03
BP	2c	0.045	0.10	1.7E-03	3.0E-02	1.3E-03
MP	2d	0.059	0.13	2.2E-03	4.5E-02	1.9E-03
EP	2d	0.019	0.04	8.3E-04	1.3E-02	5.6E-04
PP	2d	0.014	0.03	5.8E-04	1.0E-02	4.2E-04
BP	2d	0.018	0.04	6.6E-04	1.2E-02	5.0E-04

Figure 14. Rat plasma time-course simulation of exposure in the Fisher *et al.* (1999) study.

Rats were injected subcutaneously with 2 mg/kg/day BP. Only one day is shown as the clearance of parent compound is complete in less than 12 hours.



Assumptions and Uncertainties in the PBBK Modelling

The PBBK model assumes flow-limited tissue uptake for all compartments, meaning that the tissue and blood concentrations reach an equilibrium as the blood passes through the tissue. This is generally a good assumption for compounds such as parabens that do not have a high molecular weight, are not charged, and are not tightly bound to blood constituents.

Binding to plasma protein was not considered by Campbell (Campbell *et al.*, 2015), assuming that it would not impact the kinetics. Although data generated do indicate that the parabens bind to plasma proteins, it is the strength of binding and the rate of release that effect kinetics and availability for tissue uptake and metabolism. This is typically not an issue driving compound kinetics (though warfarin is a notable exception).

Two key parameters affecting dermal absorption were taken from the work of Campbell *et al.* (2015), where the values were fitted to experimental data from human exposures to BP, and human biomonitoring data for MP and PP. FracAvail represents the fraction available for absorption through the skin, directly affecting internal dose. P, the permeability parameter, drives the rate of absorption from the skin surface. Both parameters have a direct correlation with simulation output and the model is sensitive to these parameters.

The fitting of multiple parameters simultaneously to the BP data, including dermal absorption rate, fraction absorbed, renal clearance and liver conjugation rate, adds uncertainty to the PBBK simulation of the product use scenario. These parameters are not completely identifiable based on the available kinetic data, and alternate calibrations are possible that would provide similar fits to the experimental data. However, these estimates were anchored initially using *in vitro* hydrolysis measurement, which is the major route of metabolism, strengthening the confidence in the simulated results for BP.

The liver conjugation and urinary clearance rates of MP and PP were fitted to average human biomonitoring measurements of the ratio of free to total (free plus conjugates) in urine and serum. However, the serum and urine biomonitoring samples used were not matched, and exposure levels were not provided with the data. The development of these parameters assumed a constant, low-level, dermal exposure of 100 μg/day, chosen to be low enough to be in the linear range for metabolism, helping reduce the uncertainty of simulating an unknown exposure to generate ratios of free to conjugated amounts (i.e., not likely to saturate metabolism rates). Hydrolysis is the major pathway of metabolism for these parabens (see **Table 13** and **Figure 5**), primarily resulting in production of pHBA in human hepatocytes. The parameters estimated by Campbell (Campbell *et al.*, 2015) for direct conjugation of MP and PP are not small compared to the rate of hydrolysis (although it is for BP), and this is an uncertainty in the Campbell calibration in light of the missing kinetic data for verification of MP and PP simulations.

In vitro skin penetration parameters (Cross & Roberts, 2000) were adjusted for MP and PP by the ratio of fitted *in vivo* to *in vitro* skin penetration data for BP from the same study. This assumes that the *in vitro* to *in vivo* differences needed to reproduce the *in vivo* BP data are similar for the other parabens.

Plasma hydrolysis was not considered by Campbell *et al.* (2015). This is expected to be a reasonable assumption based on the NAM data collected that suggests hydrolysis in the liver exceeds that of the plasma by several orders of magnitude.

The rat SC injection dosing route has high uncertainty. The rat PBBK model was developed by Campbell based on oral exposure in adult rats. The toxicity data of Fisher *et al.* (1999)

is the result of a subcutaneous skin dose in neonatal rats. The Fisher study also suggests that there is slow release from the site of injection, though no data to support this is presented other than noting the presence of oily vehicle remaining from day to day. If the release of the paraben from the vehicle depot is indeed retarded, this could significantly impact the estimated maximum concentration of paraben in the blood. There are no rat SC kinetic data to address this uncertainty. Possible approaches to explore the magnitude of this uncertainty include analysis of the sensitivity of simulated blood concentration to the release rate of paraben from the vehicle.

The physiological and chemical-specific parameter values used for the rat simulation are shown in Appendix 4, Table 26 and Table 27. There were only rat tissue microsomal hydrolysis data in the Ozaki *et al.* (2013) study, and no data for rat cytosolic hydrolysis rates as there were for human. However, the cytosolic hydrolysis rates *in vitro* were significantly lower in the human except for skin, and the skin metabolism was turned off for the rat SC simulation as described above. Thus, the lack of cytosolic hydrolysis is assumed to have little impact on the results of the simulation.

Simulation of the product use scenario was conducted for adult consumers using parameters for an average individual. Previous evaluations of the impact of interindividual variability in pharmacokinetics on PBPK modelling of the relationship of internal dose to external exposure (Clewell & Andersen, 1996) suggest that the resulting variability in internal dose is consistent with the recommended default factor of three recommended by IPCS (IPCS, 2005). Age-dependent variations in pharmacokinetics can also be expected. However, a study of the impact of age-dependent pharmacokinetics on internal dose (Clewell *et al.*, 2004) found that, in general, predictions of average pharmacokinetic dose metrics for a chemical across life stages were within a factor of two, although larger transient variations were predicted, particularly during the neonatal period.

Appendix 6. *In vivo* legacy data

Repeat dose toxicity

Reference	OECD SIDS Initial Assessment Report for pHBA (Sponsor Country Japan) (secondary reference). Original reference Ministry of Health and Welfare (1997).
Species/Strain	Rat / Sprague Dawely
Sex	Males and females
Route of administration	Oral gavage
Exposure period	Daily dose for 45 days (males); 14 days before mating until day 3 of lactation (females)
Doses	40, 200 and 1000 mg/kg/day
Fulfilment of test guideline	Equivalent to OECD 422
Klimisch Score	2 (Reliable with Restriction)
Test substance	pHBA
NOAEL	1000 mg/kg/day (highest dose tested)
Result	pHBA induced rale and rhinorrhoea, indicative of imitation to respiratory tract irritation, and small fluctuation of blood chemistry with no changes of histopathological findings and organ weights. Changes of blood chemistry are considered not to be adverse. These effects were of no toxicological significance.
Other findings	The effects observed were considered as unspecific changes caused by the acidic nature of pHBA, which can lead to a metabolic acidosis in the blood and explains the deviations in clinical chemistry and in haematology observed after high doses of pHBA.

Reference	pHBA esters as preservatives II. Acute and Chronic Toxicity in Dogs, Rats and Mice (Matthews, Davidson, Bauer, Morrison, & Richardson, 1956).
Species/Strain	Rats / Wistar
Sex	Males and females
Route of administration	Oral (feed)
Exposure period	Up to 96 weeks
Doses	2% and 8% (equivalent to 2000 or 8000 mg/kg, respectively)
Test guideline	Pre-GLP
Klimisch Score	3 (Not Reliable)
Test substance	MP, EP, PP, BP
NOAEL	Not reliable
Result	No effects at 2%. At 8% MP, EP, and PP, decreased body weight gain and mortality was observed.
Other findings	The main toxic effect was an acute myocardial depression accompanied by hypotension, but this was transient in nature and not cumulative. No evidence of blood damage or histological changes in tissues of treated animals was noted. In concentrations below 5%, none of the esters produces any evidence of primary irritation when applied to the intact skin of man.

Developmental toxicity endpoint

Reference	Developmental Toxicity Evaluation of BP in Sprague Dawley Rats (Daston, 2004).
Species/Strain	Rats / Sprague Dawley
Sex	Female (25)
Route of administration	Oral (gavage)
Exposure period	daily on GD 6–19, inclusive (sperm positive day = GD 0)
Doses	0, 10, 100, and 1000 mg/kg/day
Test guideline	OECD 414
Klimisch Score	2 (Reliable with Restriction)
Test substance	BP
NOAEL	NOAEL: 100 mg/kg/day (maternal toxicity); 1000 mg/kg/day (foetal toxicity); LOAEL 1000 mg/kg/day

Result	Decreased maternal weight gain at highest dose tested. No differences in developmental parameters.
Other findings	Maternal weight gain was decreased at 1000 mg/kg/day, which was statistically significant during the GD 18–20 interval. Maternal food consumption was significantly decreased in the highest dose group over the dosing period (GD 6–20). There were no differences from control in any of the developmental parameters measured, including embryo/foetal viability, foetal weight, malformations, or variations.

Reproductive toxicity endpoint

Reference	Reproduction toxicity (Hoberman *et al.*, 2008).
Species/Strain	Rats / Wistar
Sex	Male
Route of administration	Oral (feed)
Exposure period	from day 22 of age for 56 days (Day 1 was the day of birth)
Doses	100, 1000, 10000 ppm for both parabens: consumed mean doses of BP = 10.9, 109 and 1088 mg/kg/day; consumed mean doses of MP = 11.2, 110 and 1141 mg/kg/day
Test guideline	Non-guideline study
Klimisch Score	1 (Reliable)
Test substance	MP, BP
NOAEL	10000 ppm in the diet (equivalent to ~1088 and 1141 mg/kg/day)
Result	No effects observed on male reproductive organs, sperm parameters, or reproductive hormones.
Other findings	Weekly measurement of serum LH, FSH and testosterone. After 56 days animals were sacrificed, sex organs were weighed and evaluated by histopathology including tubular staging of testis. Sperm evaluations were conducted including concentration and motility, daily sperm production, and morphology.

Reference	Uterotrophic assays (Hossaini, Larsen, & Larsen, 2000).
Species/Strain	Immature female mice / B6D2 F1; immature female rats / Wistar
Sex	Female
Route of administration	Mice: SC injection and oral (gavage) Rats: SC injection
Exposure period	3 consecutive days
Doses	Mice: SC injection: 5 and 100 mg/kg/day; Oral: 1, 10, 100 (MP, EP and PP) and 1000 (EP only) mg/kg/day Rats: 5 mg/kg/day (pHBA); 100, 400 and 600 mg/kg/day (BP)
Test guideline	Yes, OECD 440
Klimisch Score	2 (Reliable with Restriction)
Test substance	pHBA, MP, EP, PP, BP
NOAEL	NOAEL: 100 mg/kg/day (oral or SC) for MP, EP, PP; 400 mg/kg/day (SC) for BP in mice; LOAEL: 600 mg/kg/day (SC) for BP in rats
Result	Mice: pHBA, at sc doses of 5 or 100 mg/kg/day, and MP, EP, PP and BP at sc doses of 100 mg/kg/day did not produce any oestrogenic response in the mouse uterotrophic assay. No oestrogenic effect was observed for MP, EP and PP were administered by gavage at doses up to a top dose of 100 mg/kg/day, or for EP at 1000 mg/kg/day. No oestrogenic effect observed after dosing with a mixture of equal amounts of MP, EP and PP (total dose= 100 mg/kg/day). Rats: 5 mg/kg/day pHBA did not produce any oestrogenic effect, while BP produced a clear showed a weak oestrogenic effect response at a sc dose of 600 mg/kg/day.
Other findings	Estradiol benzoate, used as the positive control significantly increased the uterus weights of the mice, demonstrating the sensitivity of the test system.

Reference	*In vivo* oestrogen bioactivities of alkyl parabens (Lemini *et al.*, 2003).
Species/Strain	immature or ovariectomised adult mice / CD1; immature rats / Wistar
Sex	Female
Route of administration	SC injection
Exposure period	3 consecutive days

Doses	MP: 0.55, 5.5, 16.5, 55, 165 mg/kg/day; EP: 0.6, 6, 18, 60, 180 mg/kg/day; PP: 0.65, 6.5, 20, 65, 195 mg/kg/day; BP: 0.7, 7, 21, 70, 210
Test guideline	Similar to OECD 440 except use of immature rats and mice and ovariectomised mice.
Klimisch Score	2 (Reliable with Restriction)
Test substance	pHBA, MP, EP, PP, BP
NOAEL	Lowest NOELs were: 5.5 (MP), 0.6 (EP), 6.5 (PP) and 7 (BP) mg/kg/day; LOELs: 16.5 - 165 mg/kg/day for MP, 6 -180 mg/kg/day for EP, 20-65 mg/kg/day for PP, and 7 – 70 mg/kg/day for BP.
Result	Parabens increased uterine weight in immature and ovariectomised animals. pHBA was inactive in immature rats. The uterotrophic effects in immature mice had a positive correlation with the side-chain length of the ester group.
Other findings	The relative uterotrophic effect to E_2 (100) ($RUEE_2$) ranged between 34 to 91. The relative uterotrophic potencies related to E_2 (100) ($RUPE_2$) of these compounds were from 0.003 to 0.007. indicating a weak oestrogenic activity

Reference	Morphometric analysis of mice uteri treated with the preservatives MP, EP, PP and BP (Lemini *et al.*, 2004).
Species/Strain	Ovariectomised mice / CD1
Sex	Female
Route of administration	SC injection
Exposure period	3 consecutive days
Doses	MP: 55& 165 mg/kg/day; EP: 60 &180 mg/kg/day; PP: 65 & 195 mg/kg/day; BP: 70 & 210 mg/kg/day
Test guideline	OECD 440
Klimisch Score	2 (Reliable with Restriction)
Test substance	MP, EP, PP and BP
LOEL	LOEL: 55, 60, 65 and 70 mg/kg/day for MP, EP, PP and BP, respectively.
Result	Weak oestrogenic activity of SC administered parabens was observed.
Other findings	The relative uterotrophic effect to E_2 (100) ($RUEE_2$) ranged between 38 to 76. The relative uterotrophic potencies related to E_2 (100) ($RUPE_2$) of these compounds were from 0.02 to 0.009. Indicating a weak oestrogenic activity.

Reference	Some alkyl hydroxybenzoate preservatives (parabens) are oestrogenic (Routledge, Parker, Odum, Ashby, & Sumpter, 1998).
Species/Strain	immature and OVX rats / Alpk:AP.
Sex	Female
Route of administration	Oral and SC injection
Exposure period	3 consecutive days
Doses	MP: 40, 400 and 800 mg/kg/day (oral), 40 and 80 mg/kg (SC injection; BP: 40 and 400 mg/kg/day
Test guideline	OECD 440
Klimisch Score	2 (Reliable with Restriction)
Test substance	MP and BP
NOEL and LOEL	NOEL: 800 mg/kg/day for MP and BP (oral) and 40 mg/kg/day for BP (SC); LOAEL: 1200 mg/kg/day for BP (oral)
Result	Immature rats: MP administered orally or by SC injection up to 800 mg/kg/day failed to increase uterus weights in immature rats. However, BP produced a small (but statistically insignificant) increase in uterus wet and dry weights in immature rats when given orally at 800 mg to 1200 mg/kg/day. SC injection of BP significantly increased uterus wet weights between 400 and 800 mg/kg/day, and a dose of 1200 mg BP/kg/day increased uterus wet weights. The lowest dose of BP which induced a significant uterotrophic response was 200 mg/kg/day. OVX rats: SC injection of 800 mg/kg/day MP failed to increase uterus weight and vaginal cornification. SC injection of BP significantly increased uterus wet and dry weight at 1200 mg/kg/day. SC injection of BP (1000 mg/kg/day) significantly increased vaginal cornification.
Other findings	The uterotrophic response caused by BP *in vivo* was approximately 100,000 times less potent than 17b-estradiol.

Reference	Oestrogenic effects of pHBA in CD1 Mice (Lemini *et al*., 1997).
Species/Strain	Immature and adult ovariectomised mice / CD1
Sex	Female
Route of administration	SC injection
Exposure period	3 consecutive days
Doses	5, 50, and 500 mg/100 g/day
Test guideline	Similar to OECD 440
Klimisch Score	2 (Reliable with Restriction)
Test substance	pHBA
LOAEL	50 mg/kg/day
Result	Four days after treatment, PHBA produced a dose-dependent response on vaginal cornification and uterotrophic activity in both immature and adult ovariectomised mice.
Other findings	The relative uterotrophic potency of PHBA (500 mg/100 g) to E2 (1 mg/100 g) was 0.0011 in immature mice and 0.0018 in ovariectomised animals. Weak activity observed; inconsistent with *in vitro* data for pHBA.

Reference	Effect of neonatal exposure to oestrogenic compounds on development of excurrent ducts of the rat testis through puberty to adulthood (Fisher, Turner, Brown, & Sharpe, 1999).
Species/Strain	Neonatal rats / Wistar
Sex	Male
Route of administration	SC injection
Exposure period	Post-natal days 2-18
Doses	2 mg/kg/day
Test guideline	Non guideline study
Klimisch Score	3 or 4 (not assignable)
Test substance	BP
NOEL	2 mg/kg/day
Result	Testis weights of animals from 18 to 75 days of age were not affected by BP. AQP-1 immunoexpression the efferent ducts and duct morphology of rats were unaffected at any time point.
Other findings	Testis weights were decreased by 10 mg/kg GnRHa, 10 mg ethinyl estradiol, and 0.1-10 mg DES. AQP-1 immunoexpression the efferent ducts of rats was decreased by the potent oestrogens, ethinyl estradiol, and DES, but not by GnRHa. Ethinyl estradiol and DES, as well as tamoxifen, caused distension of the efferent duct lumen and an apparent reduction in height of the epithelial cells. The magnitude and duration of adverse changes induced by treatment with a range of oestrogenic compounds was broadly comparable to their oestrogenic potencies reported from *in vitro* assays.

Appendix 7. Abbreviations

AC10	concentrations associated with 10% of maximum activity
AC50	concentration of chemical associated with 50% of maximum activity
ADME	absorption, distribution, metabolism, and excretion
A_{Liver}	rate of change in the mass of the chemical in the liver
AOP	adverse outcome pathway
AR	androgen receptor
BP	butylparaben
CALUX	Chemical Activated Luciferase gene eXpression
$C_{Arterial}$	concentrations in the blood entering the liver
CES1	carboxylesterase-1
$CL_{int,\ in\ vitro}$	intrinsic clearance
Cl_{Liver}	metabolic clearance of the chemical in the liver
C_{max}	Maximum plasma concentration
C_{Venous}	concentrations in the blood leaving the liver
DART	Developmental And Reproductive Toxicology
EATS	oestrogen-, androgen-, thyroid signalling and steroidogenesis
EFSA	European Food Safety Authority
EP	ethylparaben
ER	oestrogen receptor
FAO	Food and Agriculture Organization
FDR	false discovery rate
Fup	fraction unbound in plasma
GI	gastrointestinal tract
GR	glucocorticoid receptor
H295R	angiotensin-II-responsive steroid-producing adrenocortical cell line
hTPO	human thyroid peroxidase
IL2	interleukin 2
IPCS	International Programme on Chemical Safety
iTTC	internal TTC
Kow	octanol:water partition coefficient
LOEL	lowest observed effect concentration

LogP	Logarithm of octanol/ water partition coefficient
LXR	Liver X receptor;
MCF7	breast cancer cell line
MoA	mode of action
MoE	Margin of Exposure
MoIE	margin of internal exposure
MoS	Margin of Safety
MP	methylparaben
MTORC1	mammalian target of rapamycin complex 1
NAM	New Approach Methodology; any *in silico*, *in chemico* or *in vitro* technique that may provide data or information that could support regulatory decision making
NFKB	nuclear factor kappa-light-chain-enhancer of activated B cells
NOAEL	No observed adverse effect level
NOEL	No observed effect level
PBBK	physiologically-based biokinetic modelling
pc	partition coefficients
PC10	10% of maximum oestrogenic potency
pHBA	para-hydroxybenzoic acid
PHH	primary human hepatocyte
pKa	dissociation constant
PND	post natal day
PoD	point of departure
PP	propylparaben
PPAR	peroxisome proliferator-activated receptor;
QL	liver blood flow
QSAR	Quantitative structure–activity relationship
RAR	retinoic acid receptor
REACh	Registration, Evaluation, Authorisation and Restriction of Chemicals
RfD	reference dose
RIVM	(Rijksinstituut voor Volksgezondheid en Milieu) Dutch National Institute for Public Health and the Environment
RR	risk ratio
RXR	retinoid X receptor
S9	supernatant fraction containing microsomes and cytosol

SC	subcutaneous
SCCS	Scientific Committee for Consumer Safety
STAT5	Signal transducer and activator of transcription 5
TD	toxicodynamic
TK	toxicokinetics
TNFA	tumor necrosis factor alfa
TPO	thyroid peroxidase
TR	thyroid hormone receptor
TRβ	thyroid hormone receptor β
TTC	Threshold of Toxicological Concern
TTR	transthyretin
UF	uncertainty factor
USEPA	United States Environmental Protection Agency

Appendix 8. Bibliography

Abbas S., H. Greige-Gerges, N. Karam, M. H. Piet, P. Netter, and J. Magdalou (2010), Metabolism of Parabens (4-Hydroxybenzoic Acid Esters) by Hepatic Esterases and Udp-Glucuronosyltransferases in Man. *Drug Metab Pharmacokinet,* Vol. 25, No. 6, pp. 568-77.

Arcella D, Soggiu ME, Leclercq C. (2003), Probabilistic modelling of human exposure to intense sweeteners in Italian teenagers: validation and sensitivity analysis of a probabilistic model including indicators of market share and brand loyalty. *Food Addit Contam*, Vol. 20, Suppl. 1, pp. S73-86.

Becker R. A., G. T. Ankley, S. W. Edwards, S. W. Kennedy, I. Linkov, B. Meek, M. Sachana (2015), Increasing Scientific Confidence in Adverse Outcome Pathways: Application of Tailored Bradford-Hill Considerations for Evaluating Weight of Evidence, *Regul Toxicol Pharmacol*, Vol. 72, No. 3, pp. 514-37. https://doi.org/10.1016/j.yrtph.2015.04.004.

Berggren E., A. White, G. Ouedraogo, A. Paini, A. N. Richarz, F. Y. Bois, T. Exner (2017), *Ab Initio* Chemical Safety Assessment: A Workflow Based on Exposure Considerations and Non-Animal Methods. *Comput Toxicol*, Vol. 4, pp. 31-44. https://doi.org/10.1016/j.comtox.2017.10.001.

Besselink H (2015), Validation of the (anti-)ERα CALUX bioassayU2-OS cells Transcriptional ERalpha CALUX-assay for the detection of estrogenic and anti-estrogenic chemicals for inclusion in TG455, BioDetection Systems B.V. (BDS) , the Netherlands, https://www.oecd.org/env/ehs/testing/OECD%20ERa%20Calux%20validation%20final%20v5%2011-12-15.pdf

Bessems J. G. M., A. Paini, M. Gajewska, and A. Worth (2017), The Margin of Internal Exposure (Moie) Concept for Dermal Risk Assessment Based on Oral Toxicity Data - a Case Study with Caffeine, *Toxicology*, Vol. 392, pp. 119-29. https://doi.org/10.1016/j.tox.2017.03.012.

Biesterbos W.H., *et al.* (2013), "Usage Patterns of Personal Care Products: Important Factors for Exposure Assessment." Food and Chemical Toxicology, *Food and Chemical Toxicology,* Vol. 55, pp. 8-17, https://doi.org/10.1016/j.fct.2012.11.014.

Brand W., P. E. Boon, E. V. S. Hessel, J. A. J. Meesters, M. Weda and A. G. Schuur (2018), Exposure to and toxicity of methyl-, ethyl- and propylparaben: A literature review with a focus on endocrine-disrupting properties, RIVM Report 2017-0028, National Institute for Public Health and the Environment, the Netherlands. DOI 10.21945/RIVM-2017-0028.

Brown R. P., M. D. Delp, S. L. Lindstedt, L. R. Rhomberg, and R. P. Beliles (1997), Physiological Parameter Values for Physiologically Based Pharmacokinetic Models, *Toxicol Ind Health*, Vol. 13, No. 4, pp. 407-84. https://doi.org/10.1177/074823379701300401.

Browne P., R. S. Judson, W. M. Casey, N. C. Kleinstreuer and R. S. Thomas (2015), Screening Chemicals for Estrogen Receptor Bioactivity Using a Computational Model, *Environ Sci Technol*, Vol. 49, No. 14, pp. 8804-14. https://doi.org/10.1021/acs.est.5b02641.

Campbell, J. L., M. Yoon, and H. J. Clewell (2015), A Case Study on Quantitative *in vitro* to *in vivo* Extrapolation for Environmental Esters: Methyl-, Propyl- and Butylparaben, *Toxicology*, Vol. 332, pp. 67-76. https://doi.org/10.1016/j.tox.2015.03.010.

Campbell J.L., Clewell R.A., Gentry P.R., Andersen M.E., Clewell H.J. (2012), Physiologically Based Pharmacokinetic/Toxicokinetic Modeling. *In: Reisfeld B., Mayeno A. (eds) Computational*

Toxicology. Methods in Molecular Biology (Methods and Protocols), Vol 929. Humana Press, Totowa, NJ

CIR (2017), Safety Assessment of Parabens as Used in Cosmetics; Draft Tentative Report for Panel Review. Cosmetic Ingredient Review, Washington, DC.

Clewell, H. (2005), Use of Mode of Action in Risk Assessment: Past, Present, and Future. *Regul Toxicol Pharmacol*, Vol. 42, No. 1, pp. 3-14. https://doi.org/10.1016/j.yrtph.2005.01.008.

Clewell H.J. 3rd, R.A. Clewell, and M.A. Andersen (2014) Physiologically Based Pharmacokinetic Modeling, In Principles and Methods of Toxicology, 6th Edition, edited by W.A. Hayes: CRC Press, New York, NY.

Clewell H. J., P. R. Gentry, T. R. Covington, R. Sarangapani and J. G. Teeguarden. (2004), Evaluation of the Potential Impact of Age- and Gender-Specific Pharmacokinetic Differences on Tissue Dosimetry. *Toxicol Sci*, Vol. 79, No. 2, pp. 381-93. https://doi.org/10.1093/toxsci/kfh109.

Clewell H. J., 3rd, M. E. Andersen and H. A. Barton. (2002), A Consistent Approach for the Application of Pharmacokinetic Modeling in Cancer and Noncancer Risk Assessment, *Environ Health Perspect*, Vol. 110, No. 1, pp. 85-93. https://doi.org/10.1289/ehp.0211085.

Clewell, H. J., 3rd and M. E. Andersen. (1996), Use of Physiologically Based Pharmacokinetic Modeling to Investigate Individual Versus Population Risk, *Toxicology*, Vol. 111, No. 1-3, pp. 315-29. https://doi.org/10.1016/0300-483x(96)03385-9.

Clewell, H. J., 3rd, and M. E. Andersen. (1985), Risk Assessment Extrapolations and Physiological Modeling, *Toxicol Ind Health*, Vol. 1, No. 4, pp. 111-31. https://doi.org/10.1177/074823378500100408.

Clewell, R. A., and H. J. Clewell, 3rd (2008), Development and Specification of Physiologically Based Pharmacokinetic Models for Use in Risk Assessment, *Regul Toxicol Pharmacol*, Vol. 50, No. 1, pp. 129-43. https://doi.org/10.1016/j.yrtph.2007.10.012.

Collet, B., E. Simon, S. van der Linden, N. El Abdellaoui, M. Naderman, H. Y. Man, I. Middelhof, *et al* (2019), Evaluation of a Panel of *in vitro* Methods for Assessing Thyroid Receptor B and Transthyretin Transporter Disrupting Activities, *Reproductive Toxicology*, https://doi.org/10.1016/j.reprotox.2019.05.011.

Cross, S. E., and M. S. Roberts (2000), The Effect of Occlusion on Epidermal Penetration of Parabens from a Commercial Allergy Test Ointment, Acetone and Ethanol Vehicles, *J Invest Dermatol*, Vol. 115, No. 5, pp. 914-8. https://doi.org/10.1046/j.1523-1747.2000.00151.x.

Daston, George P (2004), Developmental Toxicity Evaluation of Butylparaben in Sprague-Dawley Rats, *Birth Defects Research (Part B),* Vol. 71, No. 4, pp. 296-302. https://doi.org/10.1002/bdrb.20016.

Dearden, J., and A. Worth (2007), *In silico* Prediction of Physicochemical Properties, *JRC Scientific and Technical Reports*, EUR 23051 EN – 2007, Joint Research Centre, European Commission, Italy.

Dimitrov, S. D., R. Diderich, T. Sobanski, T. S. Pavlov, G. V. Chankov, A. S. Chapkanov, Y. H. Karakolev, *et al.* (2016), Qsar Toolbox - Workflow and Major Functionalities, *SAR QSAR Environ Res*, Vol. 27, No. 3, pp. 203-19. https://doi.org/10.1080/1062936X.2015.1136680.

ECHA, Propyl Paraben Dossier, 2018. https://echa.europa.eu/fr/registration-dossier/-/registered-dossier/13890/7/1.

Eilstein J, Grégoire S, Fabre A, A. Poupon, E. Arbey, Géniès C, Duplan H, Rothe H, Ellison C, Cubberley R, Schepky A, Klaric M, Hewitt N and Jacques-Jamin C. (2020), Use of Human Liver and Episkin™ S9 Subcellular Fractions as a Screening Assay to Compare the *in vitro* Hepatic and

Dermal Metabolism of 47 Cosmetics-Relevant Chemicals, *Journal of Applied Toxicology in press*, Vol. 40, pp 416-433. https://doi.org/10.1002/jat.3914.

Ellison, C. A., K. L. Blackburn, P. L. Carmichael, H. J. Clewell, 3rd, M. T. D. Cronin, B. Desprez, S. E. Escher , *et al.* (2019), Challenges in Working Towards an Internal Threshold of Toxicological Concern (Ittc) for Use in the Safety Assessment of Cosmetics: Discussions from the Cosmetics Europe Ittc Working Group Workshop, *Regul Toxicol Pharmacol*, Vol. 103, pp. 63-72. https://doi.org/10.1016/j.yrtph.2019.01.016.

Fagerberg, L., B. M. Hallstrom, P. Oksvold, C. Kampf, D. Djureinovic, J. Odeberg, M. Habuka , *et al.* (2014) Analysis of the Human Tissue-Specific Expression by Genome-Wide Integration of Transcriptomics and Antibody-Based Proteomics, *Mol Cell Proteomics*, Vol. 13, No. 2, pp. 397-406. https://doi.org/10.1074/mcp.M113.035600.

Fisher, J. S., K. J. Turner, D. Brown, and R. M. Sharpe (1999), Effect of Neonatal Exposure to Estrogenic Compounds on Development of the Excurrent Ducts of the Rat Testis through Puberty to Adulthood, *Environ Health Perspect*, Vol. 107, No. 5, pp. 397-405. https://doi.org/10.1289/ehp.99107397.

Fu, J., E. Pacyniak, M. G. D. Leed, M. P. Sadgrove, L. Marson, and M. Jay (2016), Interspecies Differences in the Metabolism of a Multiester Prodrug by Carboxylesterases, *J Pharm Sci*, Vol. 105, No. 2, pp. 989-95. https://doi.org/10.1002/jps.24632.

Genies, C., E. L. Jamin, L. Debrauwer, D. Zalko, E. N. Person, J. Eilstein, S. Gregoire , *et al* (2019), Comparison of the Metabolism of 10 Chemicals in Human and Pig Skin Explants, *J Appl Toxicol*, Vol. 39, No. 2, pp. 385-97. https://doi.org/10.1002/jat.3730.

Gerlowski, L. E., and R. K. Jain (1983), Physiologically Based Pharmacokinetic Modeling: Principles and Applications, *J Pharm Sci*, Vol. 72, No. 10, pp. 1103-27. https://doi.org/10.1002/jps.2600721003.

Grégoire, S., R. Cubberley, H. Duplan, J. Eilstein, D. Lange, N. Hewitt, C. Jacques-Jamin, *et al.* (2017), Solvent Solubility Testing of Cosmetics-Relevant Chemicals: Methodology and Correlation of Water Solubility to *in silico* Predictions, *Journal of Solution Chemistry*, Vol. 46, No. 7, pp. 1349-63. https://doi.org/10.1007/s10953-017-0652-5.

Hewitt, N. J., R. J. Edwards, E. Fritsche, C. Goebel, P. Aeby, J. Scheel, K. Reisinger , *et al.* (2013), Use of Human *in vitro* Skin Models for Accurate and Ethical Risk Assessment: Metabolic Considerations, *Toxicol Sci*, Vol. 133, No. 2, pp. 209-17. https://doi.org/10.1093/toxsci/kft080.

Hoberman, Alan M., David K. Schreur, Tyra Leazer, George P. Daston, Philip Carthew, Thomas Re, Linda Loretz, and Peter Mann (2008), Lack of Effect of Butylparaben and Methylparaben on the Reproductive System in Male Rats, *Birth Defects Research Part B: Developmental and Reproductive Toxicology*, Vol. 83, No. 2, pp. 123-33. https://doi.org/10.1002/bdrb.20153.

Hossaini, A., J. J. Larsen, and J. C. Larsen (2000), Lack of Oestrogenic Effects of Food Preservatives (Parabens) in Uterotrophic Assays, *Food Chem Toxicol*, Vol. 38, No. 4, pp. 319-23. https://doi.org/10.1016/s0278-6915(99)00160-x.

IPCS (2005), International Programme on Chemical, Safety Workshop on Incorporating Uncertainty Variability into Risk Assessment (2000: Berlin, Germany); Chemical-Specific Adjustment Factors for Interspecies Differences and Human Variability: Guidance Document for Use of Data in Dose/Concentration-Response Assessment, Geneva, World Health Organization. https://apps.who.int/iris/handle/10665/43294.

Jacques-Jamin, C., H. Duplan, H. Rothe, O. Vaillant, J. Eilstein, S. Gregoire, R. Cubberley, *et al.* (2017), Comparison of the Skin Penetration of 3 Metabolically Stable Chemicals Using Fresh and Frozen

Human Skin, *Skin Pharmacol Physiol*, Vol. 30, No. 5, pp. 234-45. https://doi.org/10.1159/000475472.

Janjua, N. R., H. Frederiksen, N. E. Skakkebaek, H. C. Wulf, and A. M. Andersson (2008), Urinary Excretion of Phthalates and Paraben after Repeated Whole-Body Topical Application in Humans, *Int J Androl*, Vol. 31, No. 2, pp. 118-30. https://doi.org/10.1111/j.1365-2605.2007.00841.x.

Jewell, C., P. Bennett, E. Mutch, C. Ackermann, and F. M. Williams (2007a), Inter-Individual Variability in Esterases in Human Liver, *Biochem Pharmacol*, Vol. 74, No. 6, pp. 932-9. https://doi.org/10.1016/j.bcp.2007.06.022.

Jewell, C., J. J. Prusakiewicz, C. Ackermann, N. A. Payne, G. Fate, R. Voorman, and F. M. Williams (2007b), Hydrolysis of a Series of Parabens by Skin Microsomes and Cytosol from Human and Minipigs and in Whole Skin in Short-Term Culture, *Toxicol Appl Pharmacol*, Vol. 225, No. 2, pp. 221-8. https://doi.org/10.1016/j.taap.2007.08.002.

JRC, EURL ECVAM Database on Alternative Methods to Animal Experimentation (DB-ALM), Joint Research Centre, Italy. https://ecvam-dbalm.jrc.ec.europa.eu/.

Klimisch, H. J., M. Andreae, and U. Tillmann (1997), A Systematic Approach for Evaluating the Quality of Experimental Toxicological and Ecotoxicological Data, *Regul Toxicol Pharmacol*, Vol. 25, No. 1, pp. 1-5. https://doi.org/10.1006/rtph.1996.1076.

Kolšek, K., J. Mavri, M. Sollner Dolenc, S. Gobec, and S. Turk (2014), Endocrine Disruptome--an Open Source Prediction Tool for Assessing Endocrine Disruption Potential through Nuclear Receptor Binding, *J Chem Inf Model*, Vol. 54, No. 4, pp. 1254-67. https://doi.org/10.1021/ci400649p.

Laizure, S. C., V. Herring, Z. Hu, K. Witbrodt, and R. B. Parker (2013), The Role of Human Carboxylesterases in Drug Metabolism: Have We Overlooked Their Importance?, *Pharmacotherapy*, Vol. 33, No. 2, pp. 210-22. https://doi.org/10.1002/phar.1194.

Lemini, C., A. Hernandez, R. Jaimez, Y. Franco, M. E. Avila, and A. Castell (2004), Morphometric Analysis of Mice Uteri Treated with the Preservatives Methyl, Ethyl, Propyl, and Butylparaben, *Toxicol Ind Health*, Vol. 20, No. 6-10, pp. 123-32. https://doi.org/10.1191/0748233704th202oa.

Lemini, C., R. Jaimez, M. E. Avila, Y. Franco, F. Larrea, and A. E. Lemus (2003), *In vivo* and *in vitro* Estrogen Bioactivities of Alkyl Parabens, *Toxicol Ind Health*, Vol. 19, No. 2-6, pp. 69-79. https://doi.org/10.1191/0748233703th177oa.

Lemini, C., G. Silva, C. Timossi, D. Luque, A. Valverde, M. Gonzalez-Martinez, A. Hernandez, *et al.* (1997), Estrogenic Effects of P-Hydroxybenzoic Acid in Cd1 Mice, *Environ Res*, Vol. 75, No. 2, pp. 130-4. https://doi.org/10.1006/enrs.1997.3782.

Li, B., M. Sedlacek, I. Manoharan, R. Boopathy, E. G. Duysen, P. Masson, and O. Lockridge (2005), Butyrylcholinesterase, Paraoxonase, and Albumin Esterase, but Not Carboxylesterase, Are Present in Human Plasma, *Biochem Pharmacol*, Vol. 70, No. 11, pp. 1673-84. https://doi.org/10.1016/j.bcp.2005.09.002.

Lindstedt, S. L., and P. J. Schaeffer (2002), Use of Allometry in Predicting Anatomical and Physiological Parameters of Mammals, *Lab Anim*, Vol. 36, No. 1, pp. 1-19. https://doi.org/10.1258/0023677021911731.

Lionta, E., G. Spyrou, D. K. Vassilatis, and Z. Cournia (2014), Structure-Based Virtual Screening for Drug Discovery: Principles, Applications and Recent Advances, *Curr Top Med Chem*, Vol. 14, No. 16, pp. 1923-38. https://doi.org/10.2174/1568026614666140929124445.

Manwaring, J., H. Rothe, C. Obringer, D. J. Foltz, T. R. Baker, J. A. Troutman, N. J. Hewitt, and C. Goebel (2015), Extrapolation of Systemic Bioavailability Assessing Skin Absorption and Epidermal and

Hepatic Metabolism of Aromatic Amine Hair Dyes *in vitro*, *Toxicol Appl Pharmacol*, Vol. 287, No. 2, pp. 139-48. https://doi.org/10.1016/j.taap.2015.05.016.

Matthews, Carol, Joan Davidson, Elizabeth Bauer, James L. Morrison, and Arthur P. Richardson (1956), P-Hydroxybenzoic Acid Esters as Preservatives Ii. Acute and Chronic Toxicity in Dogs, Rats, and Mice, *Journal of the American Pharmaceutical Association*, Vol. 45, No. 4, pp. 260-67. https://doi.org/10.1002/jps.3030450420.

Ministry of Health and Welfare, Japan (1997), Toxicity Testing Reports of Environmental Chemicals 5, Chemicals Investigation Promoting Council, Tokyo.

Moos, R. K., J. Angerer, G. Dierkes, T. Bruning, and H. M. Koch (2016), Metabolism and Elimination of Methyl, Iso- and N-Butyl Paraben in Human Urine after Single Oral Dosage, *Arch Toxicol*, Vol. 90, No. 11, pp. 2699-709. https://doi.org/10.1007/s00204-015-1636-0.

Munro, I. C., R. A. Ford, E. Kennepohl, and J. G. Sprenger (1996), Thresholds of Toxicological Concern Based on Structure-Activity Relationships, *Drug Metab Rev*, Vol. 28, No. 1-2, pp. 209-17. https://doi.org/10.3109/03602539608994000.

OECD (2017), Chemical Safety Assessment Workflow Based on Exposure Considerations and Non-Animal Methods. Series on Testing & Assessment No. 275, ENV/JM/MONO(2017)27, OECD, Paris. Available at http://www.oecd.org/chemicalsafety/testing/series-testing-assessment-publications-number.htm.

OECD (2016*), Test No. 455: Performance-Based Test Guideline for Stably Transfected Transactivation In Vitro Assays to Detect Estrogen Receptor Agonists and Antagonists*, OECD Guidelines for the Testing of Chemicals, Section 4, OECD Publishing, Paris, https://doi.org/10.1787/9789264265295-en.

OECD (2012), Conceptual Framework for Testing and Assessment of Endocrine Disrupters (as Revised in 2012). https://www.oecd.org/env/ehs/testing/oecdworkrelatedtoendocrinedisrupters.htm

OECD (2011), *Test No. 456: H295R Steroidogenesis Assay*, OECD Guidelines for the Testing of Chemicals, Section 4, OECD Publishing, Paris, https://doi.org/10.1787/9789264122642-en.

OECD (2004a), *Test No. 428: Skin Absorption: In Vitro Method*, OECD Guidelines for the Testing of Chemicals, Section 4, OECD Publishing, Paris, https://doi.org/10.1787/9789264071087-en.

OECD (2004b), Guidance Document for the Conduct of Skin Absorption Studies, OECD Series on Testing and Assessment, No. 28, OECD Publishing, Paris, https://doi.org/10.1787/9789264078796-en.

Ozaki, H., K. Sugihara, Y. Watanabe, C. Fujino, N. Uramaru, T. Sone, S. Ohta, and S. Kitamura. (2013), Comparative Study of the Hydrolytic Metabolism of Methyl-, Ethyl-, Propyl-, Butyl-, Heptyland Dodecylparaben by Microsomes of Various Rat and Human Tissues, *Xenobiotica*, Vol. 43, No. 12, pp. 1064-72. https://doi.org/10.3109/00498254.2013.802059.

Peyret, T., P. Poulin, and K. Krishnan (2010), A Unified Algorithm for Predicting Partition Coefficients for Pbpk Modeling of Drugs and Environmental Chemicals, *Toxicol Appl Pharmacol*, Vol. 249, No. 3, pp. 197-207. https://doi.org/10.1016/j.taap.2010.09.010.

Routledge, E. J., J. Parker, J. Odum, J. Ashby, and J. P. Sumpter (1998), Some Alkyl Hydroxy Benzoate Preservatives (Parabens) Are Estrogenic, *Toxicol Appl Pharmacol*, Vol. 153, No. 1, pp. 12-9. https://doi.org/10.1006/taap.1998.8544.

SCCS (2018), The SCCS Notes of Guidance for the Testing of Cosmetic Ingredients and Their Safety Evaluation 10th Revision, 24-25 October 2018, SCCS/1602/18, Scientific Committee on Consumer Safety, Luxembourg.

SCCS (2013), Scientific Committee on Consumer Safety (Sccs) Opinion on Parabens, 3 May 2013. Sccs/1514/13, Scientific Committee on Consumer Safety, Luxembourg.

Sivaraman, L., L. Pouliot, B. Wang, T. Brodie, M. Graziano, and M. E. McNerney (2018), Safety Assessment of Propylparaben in Juvenile Rats, *Regul Toxicol Pharmacol*, Vol. 92, pp. 370-81. https://doi.org/10.1016/j.yrtph.2017.12.009.

Sohoni, P., and J. P. Sumpter (1998), Several Environmental Oestrogens Are Also Anti-Androgens, *J Endocrinol*, Vol. 158, No. 3, pp. 327-39. https://doi.org/10.1677/joe.0.1580327.

Sonneveld, E., Riteco, J.A.C., Jansen, H.J., Pieterse, B., Brouwer, A., Schoonen, W.G., Van der Burg, B. (2006), Comparison of *in vitro* and *in vivo* screening models for androgenic and estrogenic activities. *Toxicol. Sci.*, Vol. 89, pp. 173-87.

Sonneveld, E., Jansen, H.J., Riteco, J.A.C., Brouwer, A., Van der Burg, B. (2005), Development of androgen- and estrogen-responsive bioassays, members of a panel of human cell line-based highly selective steroid responsive bioassays. *Toxicol. Sci.*, Vol. 83, pp. 136-48.

Sotoca, A. M., T. F. Bovee, W. Brand, N. Velikova, S. Boeren, A. J. Murk, J. Vervoort, and I. M. Rietjens (2010), Superinduction of Estrogen Receptor Mediated Gene Expression in Luciferase Based Reporter Gene Assays Is Mediated by a Post-Transcriptional Mechanism. *J Steroid Biochem Mol Biol*, Vol. 122, No. 4, pp. 204-11. https://doi.org/10.1016/j.jsbmb.2010.06.010.

Spicer, C.W., M.W. Holdren, R. Mukund, S.M. Gordon, and T.J. Kelly (2002), Hazardous Air Pollutant Handbook: Measurements, Properties and Fate in Ambiant Air. *CRC Press LLC,* FL, USA.

Taketani, M., M. Shii, K. Ohura, S. Ninomiya, and T. Imai (2007), Carboxylesterase in the Liver and Small Intestine of Experimental Animals and Human. *Life Sci*, Vol. 81, No. 11, pp. 924-32. https://doi.org/10.1016/j.lfs.2007.07.026.

Taylor, S., and A. Harker (2006), Modification of the Ultrafiltration Technique to Overcome Solubility and Non-Specific Binding Challenges Associated with the Measurement of Plasma Protein Binding of Corticosteroids. *J Pharm Biomed Anal*, Vol. 41, No. 1, pp. 299-303. https://doi.org/10.1016/j.jpba.2005.10.031.

Trott, O., and A. J. Olson (2010), Autodock Vina: Improving the Speed and Accuracy of Docking with a New Scoring Function, Efficient Optimization, and Multithreading. *J Comput Chem,* Vol. 31, No. 2, pp. 455-61. https://doi.org/10.1002/jcc.21334.

U.S. EPA (2011), Recommended use of body weight3/4 as the default method in derivation of the oral reference dose. Risk Assessment Forum. United States Environmental Protection Agency, Washington, D.C., EPA/100/R11/0001.

U.S. EPA (2006), Approaches for the Application of Physiologically Based Pharmacokinetic (PBPK) Models and Supporting Data in Risk Assessment (Final Report). U.S. Environmental Protection Agency, Washington, D.C., Epa/600/R-05/043f.

U.S. EPA, ToxCast Data and Information, U.S. Environmental Protection Agency, Durham, NC. https://www.epa.gov/chemical-research/toxicity-forecaster-toxcasttm-data

U.S. FDA (2018), Physiologically Based Pharmacokinetic Analyses — Format and Content Guidance for Industry, Food and Drug Administration, Silver Spring, MD, FDA-2016-D-3969.

van der Burg, B., E. B. Wedebye, D. R. Dietrich, J. Jaworska, I. Mangelsdorf, E. Paune, M. Schwarz, A. H. Piersma, and E. D. Kroese (2015), The Chemscreen Project to Design a Pragmatic Alternative Approach to Predict Reproductive Toxicity of Chemicals. *Reprod Toxicol*, Vol. 55, pp. 114-23. https://doi.org/10.1016/j.reprotox.2015.01.008.

van der Burg, B., R. Winter, M. Weimer, P. Berckmans, G. Suzuki, L. Gijsbers, A. Jonas, *et al.* (2010a), Optimization and Prevalidation of the *in vitro* Eralpha Calux Method to Test Estrogenic and Antiestrogenic Activity of Compounds. *Reprod Toxicol*, Vol. 30, No. 1, pp. 73-80. https://doi.org/10.1016/j.reprotox.2010.04.007.

van der Burg, B., R. Winter, H. Y. Man, C. Vangenechten, P. Berckmans, M. Weimer, H. Witters, and S. van der Linden (2010b), Optimization and Prevalidation of the *in vitro* Ar Calux Method to Test Androgenic and Antiandrogenic Activity of Compounds, *Reprod Toxicol*, Vol. 30, No. 1, pp. 18-24. https://doi.org/10.1016/j.reprotox.2010.04.012.

van Vugt-Lussenburg, B. M. A., R. B. van der Lee, H. Y. Man, I. Middelhof, A. Brouwer, H. Besselink, and B. van der Burg (2018), Incorporation of Metabolic Enzymes to Improve Predictivity of Reporter Gene Assay Results for Estrogenic and Anti-Androgenic Activity. *Reprod Toxicol*, Vol. 75, pp. 40-48. https://doi.org/10.1016/j.reprotox.2017.11.005.

WHO (2010), Characterization and Application of Physiologically Based Phamacokinetic Models in Risk Assessment. World Health Organization, Geneva. https://apps.who.int/iris/handle/10665/44495.

Wu, S., K. Blackburn, J. Amburgey, J. Jaworska, and T. Federle (2010), A Framework for Using Structural, Reactivity, Metabolic and Physicochemical Similarity to Evaluate the Suitability of Analogs for Sar-Based Toxicological Assessments. *Regul Toxicol Pharmacol*, Vol. 56, No. 1, pp. 67-81. https://doi.org/10.1016/j.yrtph.2009.09.006.

Ye, X., L. Y. Wong, L. T. Jia, L. L. Needham, and A. M. Calafat (2009), Stability of the Conjugated Species of Environmental Phenols and Parabens in Human Serum, *Environ Int*, Vol. 35, No. 8, pp. 1160-3. https://doi.org/10.1016/j.envint.2009.07.011.

Ye, X., A. M. Bishop, J. A. Reidy, L. L. Needham, and A. M. Calafat (2006), Parabens as Urinary Biomarkers of Exposure in Humans. *Environ Health Perspect*, Vol. 114, No. 12, pp. 1843-6. https://doi.org/10.1289/ehp.9413.

Yoon, M., J. L. Campbell, M. E. Andersen, and H. J. Clewell (2012), Quantitative *in vitro* to *in vivo* Extrapolation of Cell-Based Toxicity Assay Results. *Crit Rev Toxicol*, Vol. 42, No. 8, pp. 633-52. https://doi.org/10.3109/10408444.2012.692115.

Zuang, V., A. Dura, D. Asturiol Bofill, J. Barroso, S. Batista Leite, S. Belz, E Berggren, *et al.* (2018), EURL ECVAM status report on the development, validation and regulatory acceptance of alternative methods and approaches. EUR 29455 EN, *Publications Office of the European Union*, Luxembourg. http://dx.doi.org/10.2760/818599.