Regulatory Toxicology and Pharmacology 132 (2022) 105161

Contents lists available at ScienceDirect

Regulatory Toxicology and Pharmacology

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

Read-across and new approach methodologies applied in a 10-step framework for cosmetics safety assessment – A case study with parabens

Gladys Ouedraogo [a], Camilla Alexander-White [b], Dagmar Bury [c], Harvey J. Clewell III [d], Mark Cronin [e], Tom Cull [f], Matthew Dent [f], Bertrand Desprez [g], Ann Detroyer [a], Corie Ellison [h], Stefania Giammanco [i], Eric Hack [j], Nicola J. Hewitt [k], Gerry Kenna [l], Martina Klaric [g], Reinhard Kreiling [m], Cathy Lester [n], Catherine Mahony [n], Enrico Mombelli [o], Jorge Naciff [h], John O'Brien [i], Andreas Schepky [p], Sarah Tozer [n], Bart van der Burg [q], Barbara van Vugt-Lussenburg [q], Sharon Stuard [h], Cosmetics Europe [g,*]

[a] L'Oréal, Research & Innovation, 1 Avenue Eugène Schueller, Aulnay Sous-Bois, France
[b] MKTox & Co Ltd, 36 Fairford Crescent, Downhead Park, Milton Keynes, Buckinghamshire, MK15 9AQ, UK
[c] L'Oréal, Research & Innovation, 9 rue Pierre Dreyfus, 92110, Clichy, France
[d] Ramboll Health Sciences, 3107 Armand Street, Monroe, LA, 71201, USA
[e] School of Pharmacy and Biomolecular Sciences, Liverpool John Moores University, Byrom Street, Liverpool, L3 AF, UK
[f] Unilever, Safety & Environmental Assurance Centre, Colworth House, Sharnbrook, Bedfordshire, MK44 1ET, UK
[g] Cosmetics Europe, 40 Avenue Hermann-Debroux, Brussels, Belgium
[h] The Procter & Gamble Company, Cincinnati, OH, 45040, USA
[i] Creme Global, 4th Floor, The Design Tower, Trinity Technology & Enterprise Campus, Grand Canal Quay, Dublin 2, Ireland
[j] ScitoVation, Research Triangle Park, Durham, NC, USA
[k] SWS, Wingertstrasse 25, 64390, Erzhausen, Germany
[l] 2 Farmfield Drive, Macclesfield, Cheshire, SK10 2TJ, UK
[m] Clariant Produkte (Deutschland) GmbH, Am Unisys-Park 1, 65843, Sulzbach, Germany
[n] Procter & Gamble Technical Centres Ltd, Reading, RG2 0RX, UK
[o] Institut National de l'Environnement Industriel et des Risques (INERIS), Parc ALATA BP2, 60550, Verneuil en Halatte, France
[p] Beiersdorf AG, Global Toxicology, Unnastrasse 48, 20253, Hamburg, Germany
[q] BDS, Science Park 406, 1098XH, Amsterdam, the Netherlands

ARTICLE INFO

Handling Editor: Dr. Lesa Aylward

Keywords:
Butylparaben
Propylparaben
Ethylparaben
Methylparaben
Read-across (RAX)
New approach methodologies (NAM)
Next generation risk assessment (NGRA)
Systemic toxicity
Physiologically based kinetic (PBK) modelling

ABSTRACT

Parabens are esters of para-hydroxybenzoic acid that have been used as preservatives in many types of products for decades including agrochemicals, pharmaceuticals, food and cosmetics. This illustrative case study with propylparaben (PP) demonstrates a 10-step read-across (RAX) framework in practice. It aims at establishing a proof-of-concept for the value added by new approach methodologies (NAMs) in read-across (RAX) for use in a next-generation risk assessment (NGRA) in order to assess consumer safety after exposure to PP-containing cosmetics. In addition to structural and physico-chemical properties, *in silico* information, toxicogenomics, *in vitro* toxicodynamic, toxicokinetic data from PBK models, and bioactivity data are used to provide evidence of the chemical and biological similarity of PP and analogues and to establish potency trends for observed effects *in vitro*. The chemical category under consideration is short (C1–C4) linear chain n-alkyl parabens: methylparaben, ethylparaben, propylparaben and butylparaben. The goal of this case study is to illustrate how a practical framework for RAX can be used to fill a hypothetical data gap for reproductive toxicity of the target chemical PP.

1. Introduction

Parabens are esters of para-hydroxybenzoic acid (pHBA) and are used widely 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 and to prevent microbiological spoilage. Short-chain n-alkyl parabens have

* Corresponding author. Cosmetics Europe, Avenue Hermann Debroux 40, 1160, Brussels, Belgium.
E-mail address: cosmeticseurope@cosmeticseurope.eu (Cosmetics Europe).

https://doi.org/10.1016/j.yrtph.2022.105161
Received 25 November 2020; Received in revised form 23 January 2022; Accepted 17 March 2022
Available online 1 May 2022

been used in cosmetic products for decades. Consumers may use different types of cosmetic products daily that contain propylparaben (PP). Therefore, an estimate of total aggregate external exposure to PP from cosmetic products is considered, and further refined by using internal dose metrics from physiologically-based kinetic (PBK) modelling as relevant to humans. To perform a NGRA, a point of departure (POD) needs to be defined on the basis of hazard data with which to compare the exposure estimate. This case study shows how RAX can be used and how a point of departure is defined also as an internal dose metric, in order to derive a margin of internal exposure (MoIE). The current case study assumes, hypothetically, that there are no *in vivo* reproductive toxicity data available for PP. Due to the implementation of the ban on animal testing for cosmetics products that came into force in the EU in March 2013 it is generally not possible to generate any new *in vivo* animal data to fill data gaps or refine knowledge for cosmetic ingredients marketed in the European Union (EU). Therefore, new ways must be found to provide evidence for the safety of cosmetics ingredients without animal testing.

The 10-step RAX framework is followed, as described in our accompanying paper (Alexander-White et al., 2022) and is based on the outcome of the EU SEURAT-1 project (Berggren et al., 2017), to show how the safety of PP as used in cosmetics can adequately be assessed, without the need to generate new animal testing data. In this example, *in vivo* data have been used to draw upon existing information on systemic toxicity, i.e. data that have been generated prior to the March 2013 ban on animal testing in the EU. A tiered approach (Tiers 0, 1 and 2) is taken in the 10-step RAX framework where all existing information on the target chemical, PP, is reviewed and source analogues selected based on properties related to chemical structure as well as a hypothesised mechanism of action and an understanding of systemic exposure. As a result of performing chemical similarity profiling, analogue searches and hypothesis generation in Tier 0, three related short (C1–C4) linear chain n-alkyl parabens come under consideration as the best category based on PP being part of a homologous series in this case study, namely methylparaben (MP), ethylparaben (EP), PP and butylparaben (BP). In Tiers 1 and 2, we show as a proof of concept how NAMs, including toxicogenomics, bioavailability, kinetic data, and other biological assay data, can be integrated to consider the biological similarity, substantiate the mechanism of action and assess relative potency differences of the chemicals in the parabens category, and how this evidence can be applied in a NGRA for low toxicity chemicals.

The goal of this case study is therefore, primarily to demonstrate how NAMs can be used to support RAX, integrating both toxicodynamic and toxicokinetics data. Specifically, the case study highlights employing PBK 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.

Abbreviations:

Butylparaben (BP)	
propylparaben (PP)	
ethylparaben (EP)	
methylparaben (MP)	
pHBA	para-hydroxybenzoic acid
pHHA	para-hydroxyhippuric acid
ADME	absorption, distribution, metabolism, excretion
CPR	Cosmetic Products Regulation
CSR	Cosmetic Safety Report
EU	European Union
IP	intraperitoneal
IV	intravenous
MACCS	Molecular ACCess System
MOA	mode of action
MOIE	Margin of Internal Exposure
MOS	Margin of Safety
NAM	New Approach Methodologies
NGRA	Next Generation Risk Assessment
PBK	physiologically-based kinetic
POD	point of departure
PBS	phosphate-buffered saline;
RAX	Read-Across using new approach methods
RPF	relative potency factor
SCCS	Scientific Committee on Consumer Safety
SEURAT	Safety Evaluation Ultimately Replacing Animal Testing
SMILES	Simplified Molecular Input Line Entry Specification
TTC	Threshold of Toxicological Concern

2. Applying the 10-step RAX framework in a NGRA for propylparaben

This NGRA approach follows the recommendations of the SEURAT-1 project (Berggren et al., 2017); the tiered 10-step RAX framework (Alexander-White et al., 2022) that is applied in this case study for performing a NGRA for PP is shown in Fig. 1. This paper walks through the framework and how it is applied in practice to reach a human safety decision for the safety of PP in cosmetic products, with the focus on reproductive toxicity endpoints.

2.1. Problem formulation

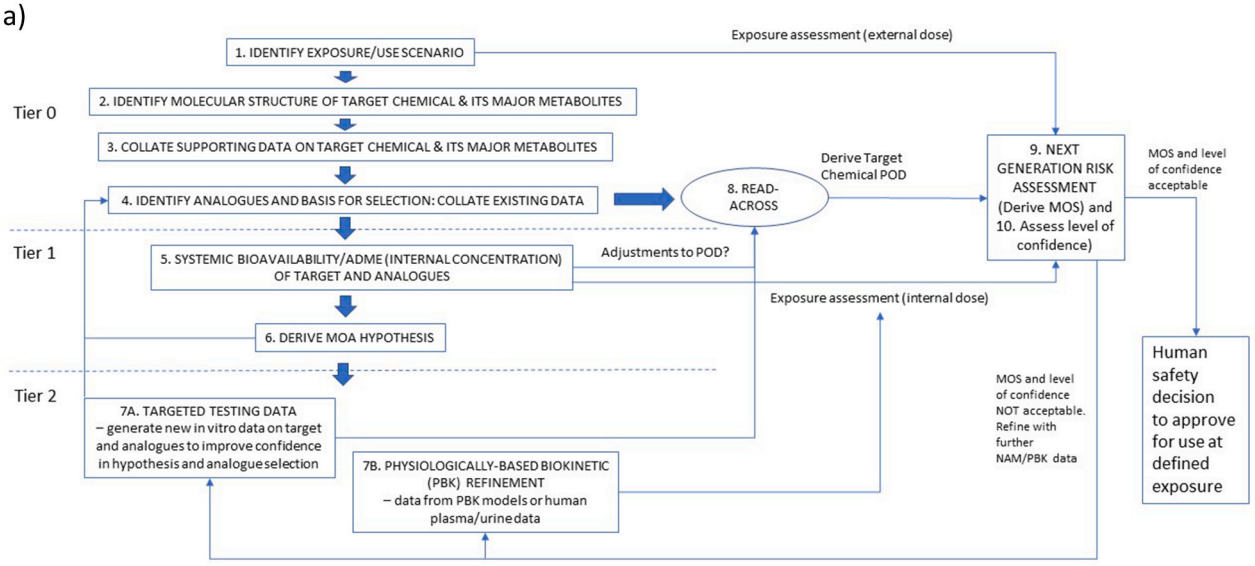
For all NGRA it is important to begin with clear problem formulation. In this case, human safety of the target substance PP has to be assured despite the (hypothetical) lack of *in vivo* data on reproductive/developmental toxicity, as this is considered the pivotal endpoint on the basis of an assumed endocrine-related MOA. It has to be decided what represents an acceptable 'safe' concentration in a product, without the need for new animal testing. A tiered approach is followed to assess dermal exposure to PP in cosmetics, as would be applied in the SCCS Notes of Guidance (2018). The 10-step RAX framework is adopted to perform a NGRA based on NAMs for reproductive endpoints. Tier 0 utilizes the threshold of toxicological concern (TTC) (Munro et al., 1996), and if consumer exposure estimates exceed the TTC level for PP, an appropriate reproductive/developmental toxicity POD is needed for safety assessment. In that case, subsequent tiers will be followed to conduct a RAX informed by NAM to address the data gap. NAMs (PBK modelling) will also be used to better understand the systemic exposure to parabens. At this point in time, TTC involves external exposure doses, but work is underway to consider a potential future internal dose TTC approach (Ellison et al., 2019).

3. *Tier 0* - steps 1 to 4 of next generation read-across (NGRA)

The first tier of the framework involves defining exposure for PP, searching for existing data and identifying analogue(s) for a RAX hypothesis. At the start of the process, as PP is an existing cosmetic ingredient, a deterministic or probabilistic exposure estimate can be provided based on known cosmetic product use.

Step 1) Identify Exposure/Use scenario for PP in cosmetic products

a)



b)

Tier 0

Step 1: Identify exposure/use scenarios for target chemical

Step 2: Identify molecular structure of target chemical

Step 3: Collate supporting data on target chemical and define data gap(s)

Step 4: Analogue(s) a) Identify, b) collate existing data, c) determine similarity hypothesis

End Tier 0 → Potential to move to Steps 8-10 if data are sufficient

Tier1

Step 5: Systemic bioavailability/ADME of target chemical and analogues

Step 6: Supporting a Similar Mode/Mechanism of Action (MoA) hypothesis

End Tier 1 → Potential to move to Steps 8-10 if data are sufficient

Tier 2

Step 7: a) Perform targeted testing to strengthen hypotheses and/or b) Biokinetic refinements of target chemical and analogues

The Assessment

Step 8: Performing a read-across (RAX) to derive a point of departure (POD)

Step 9: Performing a margin of safety (MOS) evaluation

Step 10 Assessing the level of confidence for establishing if the MOS is acceptable

Fig. 1. A tiered 10-step framework (as in Alexander-White et al. (2022)) to enable a human safety decision to be made using NAMs and RAX, which in (a) diagrammatically builds on the SEURAT 1 workflow (Berggren et al., 2017) to perform a next generation risk assessment (NGRA) without new animal data; the steps are tabulated in (b).

The initial step for this case study is to derive an external dermal exposure dose metric for PP in cosmetics products. PP is used in cosmetic products at a maximum concentration of 0.18% (N.B. this is a current regulatory maximum in the EU for PP, and has been used in this case study for illustration, as it has been defined by previous regulatory assessment in the EU for parabens; SCCS 2013).

A simple deterministic consumer exposure estimate for PP in adults includes the maximum allowed use concentration of 0.18% PP, and a

maximum estimated daily exposure level for different cosmetic products of 17.4 g/day per SCCS Notes of Guidance (SCCS, 2018). This aggregate exposure scenario for dermally applied products (based upon the data in SCCS 2018) results in an external dose estimate of 0.53 mg PP/kg bw/day. This is a theoretical worst-case scenario as it assumes all cosmetic products contain the maximum PP level and are used by all persons at a high amount per use, at a high frequency per day, simultaneously, which is clearly not a realistic scenario (SCCS, 2018). Nonetheless it is a simple set of conservative assumptions to begin, and taking this approach may in some cases lead to an acceptable risk assessment outcome.

As comprehensive survey data across the EU Cosmetics Industry in 2016 were available on real use levels and occurrence data of PP in cosmetic products, a higher tier probabilistic external dose exposure assessment was performed on the basis of real European consumers' habits & practices (H&P) data in the Creme Care and Exposure model (https://www.cremeglobal.com/; and as exemplified in Tozer et al., 2015 for zinc pyrithione and (in Tozer et al., 2019) for vitamin A exposure. Moving to a probabilistic and subject-oriented model can provide refinement of the estimates of exposure. This probabilistic modelling also allowed for the use of statistical distributions to characterize substance concentrations and the use of product occurrence data to account for the presence of chemicals in some, but not all products.

Four scenarios were considered in Creme Care and Exposure modelling:

a. Paraben always present, max concentration as per regulation
b. Paraben always present, concentration at current use range according to Cosmetic Europe Product Preservation Survey (2016)
c. Using paraben occurrence data according to Mintel GNPD, max concentration in regulation
d. Using paraben occurrence data according to Mintel GNPD, concentration at current use range according to Cosmetic Europe Product Preservation Survey (2016).

The external dermal exposure dose metrics for PP as calculated in the Creme Care and Exposure Model for each scenario (p95) were (in mg/kg/day) a) 0.154, b) 0.057, c) 0.084 and d) 0.014. Of these, scenario (d) represents the most realistic exposure scenario for PP exposure through use of cosmetic products.

Step 2) Identify molecular structure of PP and its major metabolite(s)

The target substance PP is a white crystalline solid at room temperature and has the chemical structure as shown in Fig. 2. PP is the propyl (C3 n-alklyl) ester of pHBA. It is stable in air and does not hydrolyse in hot or cold water or in acidic conditions.

The structure shows that PP can be hydrolysed to propanol and pHBA.

Step 3) Collate supporting data for PP and its major metabolite(s) With the best possible exposure estimate (from Step 1) and a knowledge of the chemical structure of the target substance and its major metabolite(s) (from Step 2) to address whether there are any known toxicity alerts according to Cramer classification, it is possible to exit the framework, if exposure is less than a TTC (Kroes et al., 2007; Yang et al., 2017; EFSA, 2019; Mahony et al., 2020). Using a deterministic dermal exposure metric for PP in cosmetics products of 0.53 mg/kg/day (see Step 1 above), a TTC approach is not possible as this estimated intake is higher than the threshold level for Cramer class I (stated as 0.042 mg/kg/day (Yang et al., 2017) or 0.03 mg/kg/day (EFSA, 2019)) to which PP is allocated due to its simple chemical structure with no alerting functional groups and simple ester hydrolysis leading to innocuous end products (propanol and pHBA) suggesting a low order of general toxicity. The output from the Creme Care and Exposure probabilistic model for PP from scenario (d) in step 1 yields an exposure for PP of 0.014 mg/kg/day, which would enable the use of the TTC approach at this point, as this exposure is lower than the required TTC threshold of 0.03 mg/kg/day.

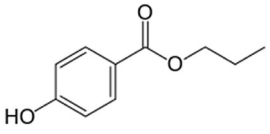


Fig. 2. Chemical structure of propyl paraben ($C_{10}H_{12}O_3$; CAS RN 94-13-3: SMILES CCCOC(=O)C1=CC=C(C=C1)O).

However, to illustrate the 10-step RAX framework and show how even more assurance can be given for the safety of PP using exposure based NGRA, we progress on to Step 3 of a RAX. In Step 3, we collate all existing data on the target substance, including physico-chemical parameters, relevant toxicology information and existing assay data etc.

The physico-chemical properties of PP are as described below in Table 2. For illustration in this case study, as to how chemical similarity as a first step drives analogue selection, we have assumed there is no *in vivo* toxicity data for PP at this point in the process.

Using the chemistry information, a search for similar analogues is performed.

Step 4) Identify analogue(s) and basis for selection (a):collate existing data (b)

a) Identify analogues and basis for selection

Suitable analogues were identified using the expert-judgement based method of Wu et al. (2010) which relies on consideration of similarity in structure, metabolism, reactivity and physical chemical properties. Substructure searching is performed using a defined molecular scaffold with required functional groups. For the parabens, analogues must possess a phenyl ring with a hydroxyl group and a carboxylic acid group esterified to an aliphatic alcohol of variable chain length. Tanimoto comparisons of molecular fingerprints also may be considered and may be used in combination with substructure searching for identifying potential analogues, however, structural similarity scores alone should not be used to justify the suitability of an analogue for RAX and must be combined with considerations of metabolism, reactivity and physical chemical properties (Lester et al., 2018).

The three short-chain parabens MP, EP and BP are identified as potential source analogues for PP and are displayed in Table 1. Structural similarity scores are included in the table and were calculated using a Tanimoto algorithm for comparing molecular fingerprints generated using the proprietary 960 structural keys from Biovia Corp (https://www.3ds.com/products-services/biovia/). Structural differences between the three analogues and PP include differences in the alcohol chain length which is C1 for MP, C2 for EP, C3 for PP and C4 for BP. The calculated similarity scores comparing the structures of MP, EP and BP with PP are 0.81, 0.93 and 0.94, respectively and are consistent with small differences in structure.

Similarity in biotransformation pathways must be considered when determining the suitability of an analogue for RAX. The predominant metabolic pathway for the short-chain parabens listed in Table 1 is known to be hydrolysis of the ester bond to form the common primary metabolite pHBA, its glycine conjugate p-hydroxyhippuric acid (pHHA) and the corresponding alcohol (CIR 2008; Shin et al., 2019; Géniès et al., 2019). Studies in humans have shown that parabens also can be excreted as glucuronide and sulfate conjugates (Soni et al., 2005).

Analogue suitability also depends on the values of physicochemical properties relative to those for the target chemical. Physicochemical properties can affect bioavailability and consequently biological

Table 1
Chemical structures, molecular weight and Tanimoto similarity of category members in the homologous series of parabens.

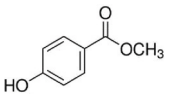
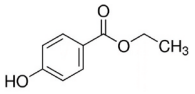
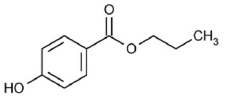
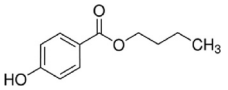
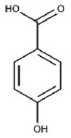
Target and Source Chemicals	Chemical Name	CAS No.	Molecular Formula	Molecular Weight	Chemical Structure	Similarity (Tanimoto coefficient)
Source 1	Benzoic acid, -4-hydroxy-, methyl ester (Methylparaben; MP)	99-76–3	$C_8H_8O_3$	152		0.81
Source 2	Benzoic acid, -4-hydroxy-, ethyl ester (Ethylparaben; EP)	120-47–8	$^{C}9^{H}10^{O}3$	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_7H6O_3	138		Not applicable

Table 2
Comparison of physico-chemical properties of the target substance propyl paraben and analogues methyl paraben, ethyl paraben, butyl paraben. Measured or a: predicted from EPA EPI Suite version 4.1; b: predicted from OECD QSAR Toolbox v4.2.

Parameter	MP	EP	PP	BP
Molecular weight[a]	152.15 (a)	166.17 (a)	180.2 (a)	194.23 (a)
Melting Point (°C)	131 °C (b)	117 °C (b)	97 °C (b)	68.5 °C
Volatility (mmHg at 25 °C)[a]	0.000855 (a)	0.0000929 (a)	0.000307 (a)	0.000251 (a)
LogP	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 at 25 °C	8.34–8.87 (b)	8.18–8.9 (b)	7.91–8.87 (b)	8.34
Aqueous solubility (mg/L)	2500 at 25 °C (a)	885 at 25 °C (a); 885 (b)	500 at 25 °C (a, b)	207 at 20 °C

[a] Calculated values.

responses observed *in vitro* or *in vivo*. The key physicochemical properties which could affect bioavailability (Lipinski et al., 2001) of the four parabens are listed in Table 2.

In review of the results in Table 2, it can be seen that with increasing length of nonpolar side chains the logP increases steadily and water solubility decreases. Although in the case of propyl versus the three paraben analogues, the differences are within admitted experimental variation for solubility properties (Dearden and Worth 2007). These properties can impact the relative bioavailability of the parabens, particularly when considering the bioavailability after dermal application. In general, lipophilic substances with a $logP_{OW}$ above 3 show a lower skin penetration rate than more hydrophilic substances with a $logP_{OW}$ between −1 and 3 due to deposition in the lipophilic matrix of the skin such as the dermis (Danish EPA, 2009). The predicted volatility of all four parabens is very low. From the pKa values, the acid/basic behaviour of all four parabens is essentially the same. As they all have a pKa of ~8 associated to an acidic function, these short linear chain parabens would be expected to show similar patterns of bioavailability.

b) Collation of existing data for the selected analogues

3.1. Legacy in vivo data

One approach to substantiate analogue(s) for the purposes of using a suitable human-relevant toxicological POD in a risk assessment is to source the available reproductive/developmental toxicity data, and assess the quality of the study and the confidence in the POD.

The key *in vivo* reproductive/developmental toxicity studies on the three parabens (MP, EP and BP) and pHBA source chemicals (conducted prior to 2013) were reviewed and are presented in Table 3.

In addition, *in vivo* screening studies such as the uterotrophic assay are summarised in Table 4. It has to be noted that the uterotrophic assay in either immature or ovariectomised rodents is a short-term screening assay on biological (oestrogenic) activity of the respective substance. The measured endpoint is an increase in uterus weight which presents no evidence of an endocrine-mediated adverse effect.

Overall, the valid (Klimisch score 1 and 2; Klimisch et al., 1997) *in vivo* reproductive/developmental toxicity data in Table 3 demonstrate no relevant adverse reproductive effects for MP, EP and BP at oral (diet, gavage) doses up to 1000 mg/kg/day. A POD of 1000 mg/kg/day has been used for MP and EP in regulatory risk assessment for the past two decades, and there is no concern over the safety of these paraben analogues. The results from the studies by Oishi (2001, 2002) are not regarded as valid (Klimisch score 3), as they were derived from non-guideline and non-GLP studies where the documentation was neither sufficient nor the effects regarded as biologically plausible. In addition, other working groups using the same test protocol and rat strain as those used in the studies by Oishi failed to reproduce these effects at a dose up to 1000 mg/kg bw/day although a larger number of animals and additional reproductive endpoints were included (Hoberman et al., 2008). Also, no adverse effect of BP was reported after the rats received a SC injection of 2 mg/kg bw (Fisher et al., 1999). The POD for BP of 2 mg/kg/day is derived from a sub-cutaneous (SC) route of exposure. It has to be emphasised that the SC route of exposure circumvents the skin barrier. The skin is known to metabolise parabens effectively by skin esterases (Williams, 2008). SC dosing may considerably increase the internal bioavailability of parent paraben compared to dermal exposure (Aubert et al., 2012). As dermal application is the major exposure route for cosmetic products, dermal absorption and metabolism need to be considered for the safety assessment of PP.

As this was the only dose tested in this chosen pivotal study for this case illustration, it is an extremely conservative POD as the NOEL may

Table 3
Legacy reproductive and developmental toxicity data for source chemicals methylparaben, ethylparaben and butylparaben and primary metabolite pHBA, and with suitable quantitative data to define a point of departure, either as a no observed (adverse) effect level (NOAEL/NOEL) or a lowest observed adverse effect level (LOAEL).

Study Details, Klimisch Score	Results	NOAEL mg/kg/day	LOAEL mg/kg/day	Reference
Methyl paraben				
Reproductive toxicity studies – male reproduction				
Non-guideline bespoke study investigating effects on male reproduction. Male rats fed diets containing 10000 ppm MP 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. GLP; Klimisch 1	No effects observed on male reproductive organs or parameters up to the top dose (1000 ppm).	10000 ppm in the diet (equivalent to **1088 mg/kg/day**) (Top dose)	N/A	Hoberman et al. (2008)
MP was administered to groups of eight 3-week-old male Wistar rats at doses of 0%, 0.1%, and 1.0% each in the diet for eight weeks, corresponding to average intakes of 103 and 1030 mg MP/kg bw/day. Non-guideline, nonGLP; Klimisch 3.	No effects were observed on weights of the reproductive organs, on sperm counts in the testes and epididymides, and on the morphological examinations of spermatogonia, spermatocytes, round spermatids and elongated spermatids. In addition, serum concentrations of testosterone, LH and FSH were not affected.	**1030**	N/A	Oishi (2004)
Ethyl paraben				
Reproductive toxicity studies – male reproduction				
EP was administered to groups of eight 3-week-old male Wistar rats at doses of 0.00%, 0.1%, and 1.0% each in the diet for eight weeks, corresponding to average intakes of 103 and 1030 mg methyl paraben/kg bw/day and 103 and 1043 mg ethyl paraben/kg bw/day, respectively. Non-guideline, nonGLP; Klimisch 3	No effects were observed on weights of the reproductive organs, on sperm counts in the testes and epididymides, and on the morphological examinations of spermatogonia, spermatocytes, round spermatids and elongated spermatids. In addition, serum concentrations of testosterone, LH and FSH were not affected.	**1043**	N/A	Oishi (2004)
Propyl Paraben – theoretical data gap (Target Chemical)				
Butyl paraben				
Reproductive toxicology – male reproductive effects				
Non-guideline bespoke study investigating effects on male reproduction. Male rats fed diets containing 10000 ppm MP 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. GLP; Klimisch 1	No effects observed on male reproductive organs or parameters up to the top dose (1000 ppm).	10000 ppm in the diet (equivalent to **1088 mg/kg/day**) (Top dose)	N/A	Hoberman et al. (2008)
Developmental effects				
OECD 414 Prenatal Developmental Toxicity Study in Sprague Dawley rats at oral (gavage) doses of 10, 100 and 1000 mg/kg/day. GLP; Klimisch 1	Decreased maternal weight gain at highest dose tested. No differences in developmental parameters	100 (maternal) 1000 (fetal)	1000 (maternal)N/A (fetal)	Daston (2004)
BP was administered to groups of eight 3-week-old male Wistar rats at doses of 0.00%, 0.01%, 0.10% and 1.00% in the diet for eight weeks, corresponding to average butyl paraben intakes of 10, 100 and 1000 mg/kg/day. Non-guideline, nonGLP, study refuted by Hoberman et al. (2008); Klimisch 3	The weights of the epididymides were significantly decreased in the mid- and high-dose groups. The cauda epididymal sperm reserve of all treated groups was decreased. The sperm count of the high dose group was 58.2% of the control value. The daily sperm production in the testis was also significantly lower in all treated groups. Serum testosterone was significantly decreased at the mid and high doses.	N/A	10 mg/kg/day	Oishi (2001)
BP was administered to groups of eight 4-week-old male Crj:CD-1 mice at doses of 0.00%, 0.01%, 0.10% and 1.00% in the diet for 10 weeks, corresponding to average butyl paraben intakes of 14.4, 146 and 1504 mg/kg bw/day, respectively. Non-guideline, nonGLP; Klimisch 3	The weights of the epididymides were significantly increased in the high-dose group. A dose-dependent decrease of both round and elongated spermatid counts was observed in the seminiferous tubules. The number of spermatogonia and spermatocytes were not different from the controls. Serum testosterone was significantly decreased at the highest dose	14.4 mg/kg/day	146 mg/kg/day	Oishi (2002)
Single dose. Neonatal Wistar rats were administered by SC injection with 2 mg/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. Non-guideline, nonGLP; Klimisch 3	No alteration in testis weights when compared to control animals at day 18.	No effects at the only tested dose of 2 mg/kg/day, the NOEL/NOAEL cannot be determined as only one dose tested	N/A	Fisher et al. (1999)

(*continued on next page*)

Table 3 (*continued*)

Study Details, Klimisch Score	Results	NOAEL mg/kg/day	LOAEL mg/kg/day	Reference
Non-guideline study where pregnant Sprague-Dawley rats were injected subcutaneously with 100 or 200 mg/kg of BP from gestation day (GD) 6 to postnatal day (PND) 20.	In the group exposed to 200 mg/kg of BP, the proportion of pups born alive and the proportion of pups surviving to weaning were decreased. The body weights of female offspring were significantly decreased at PND 49. The weights of testes, seminal vesicles and prostate glands were significantly decreased in rats exposed to 100 mg/kg of BP on PND 49. In contrast, the weights of female reproductive organs were not affected by BP. The sperm count and the sperm motile activity in the epididymis were significantly decreased at doses of 100 and 200 mg/kg of BP.			Kang et al. (2001)
Para-hydroxybenzoic acid (primary metabolite)				
OECD combined repeated dose and reproductive/developmental toxicity screening test. 4-Hydroxybenzoic acid was administered by gavage at doses of 40, 200 and 1000 mg/kg for 45 days in males and from 14 days before mating to day 3 of lactation in females. Klimisch 1	No adverse effects on copulation, fertility, maintenance of pregnancy, parturition and lactation, as well as viability, sex ratio, body weights and morphological appearance of pups at all treated groups.	1000 mg/kg/day (parent and offspring)	N/A	MHW, Japan (1997)
Oral toxicity study (day 11 of gestation) was performed in pregnant Sprague-Dawley rats at single doses of 333, 667, 1000 mg/kg. EPA; Klimisch 2	No maternal toxicity, including death and change in body weight gain at 24 and 72 h after treatment. In addition, no developmental toxicity was observed, including change in litter size, pup weight, and total litter weight at 1 and 6 days after birth, and overt malformation.	1000 mg/kg/day	N/A	Kavlock (1990)

Table 4
Legacy uterotrophic assay data for target and source chemicals methyl paraben, ethyl paraben and butyl paraben and the primary metabolite pHBA (Klimisch score 1 or 2). ND = not determined.

Study details/Klimisch score	Methyl Paraben	Ethyl Paraben	Butyl Paraben	Para-Hydroxy Benzoic Acid	Reference
Appears compliant with OECD 440 Uterotrophic Study. ovariectomised CD1 mice, SC doses for 3 days. NonGLP, Klimisch 2	Weak oestrogenic activity observed at 55 and 165 mg/kg/day	Weak oestrogenic activity observed at 60 and 180 mg/kg/day	Weak oestrogenic activity observed at 70 and 210 mg/kg/day	ND	Lemini et al. (2004)
Appears compliant with OECD 440, immature rats and mice and ovariectomised mice. SCdoses for 3 days. NonGLP, Klimisch 2	Weak oestrogenic activity observed at 16.5–165 mg/kg/day, no activity at 5.5 mg/kg/day	Weak oestrogenic activity observed at 6–180 mg/kg/day, no activity at 0.6 mg/kg/day	Weak oestrogenic activity observed at 7–70 mg/kg/day, no activity at 7 mg/kg/day	ND	Lemini et al. (2003)
Appears compliant with OECD 440 Uterotrophic Study. immature female B6D2F1 mice, oral and SC doses for 3 days. NonGLP, Klimisch 2	No effects of MP at any dose tested NOEL 100 mg/kg/day (oral and SC) (top dose)	No effects of EP at any dose tested NOEL 100 (SC) NOEL 1000 (oral) (top dose)	Weak oestrogenic activity observed NOEL 400 (SC) LOAEL 600 (SC)	No effects of pHBA at any dose tested NOEL 100 mg/kg/day (oral and SC) (top dose)	Hossaini et al. (2000)
Appears compliant with OECD 440 Uterotrophic Study. MP administered orally and SC (up to 800 mg/kg/day) and BP orally (up to 800 mg/kg/day) and SC (up to 1200 mg/kg/day) to immature female Alpk:AP rats. NonGLP, Klimisch 2.	No increase in uterine weights at any dose up to 800 mg/kg/day (oral and SC)	ND	Weak oestrogenic activity observed at 1200 mg/kg/day (oral), no activity at 800 mg/kg/day (oral) and 40 mg/kg/day (SC), approximately 100,000 times less potent than 17 beta-estradiol	ND	Routledge et al. (1998)

(and indeed has been proven to) be much higher than this. It was nevertheless selected as a NOEL by the Scientific Committee for Consumer Safety in the opinion on parabens (SCCS, 2013) in the absence of further robust information at that time. The SCCS acknowledged that the choice of this POD was very conservative and unusual in terms of the SCCS Notes of Guidance and general principles of risk assessment, thus, this low and not well established POD was considered provisional at the time, aiming at protecting the consumers in a very conservative manner until further data became available.

The uterotrophic assay data (Table 4) show that the parabens are broadly similar in terms of the weak biological activity, if any, determined in this kind of *in vivo* screening assay. The metabolite pHBA showed no effects in reproductive/developmental toxicity studies nor in *in vivo* screening assays for endocrine activity such as the uterotrophic assay, and therefore it was considered that any adverse effects observed would not be due to this shared main metabolite (Kavlock, 1990; Ministry of Health and Welfare Japan, 1997; OECD, 1999; Hossaini et al., 2000).

The findings of the available *in vivo* studies on parabens and pHBA prompted the consideration of whether existing NAM data (*in silico* profiling and *in vitro* data) particularly related to endocrine activity, focused on e.g. Oestrogenic, Androgenic, Thyroidal, Steroidogenic (EATS), could help to provide mechanistic hypotheses in principle for this RAX category and help in affirming analogue identification. Comparing the mechanistic profiles and potencies of the target and source compounds in relevant NAM assays would help in defining the POD and how it can be applied in the risk assessment.

3.2. Existing in silico profiling data for parabens with a focus on reproductive toxicity and related endocrine activity

In addition to the evaluation of physicochemical properties and available *in vivo* data, an analysis of *in silico* data with focus on the respective RAX endpoint is important in determining the similarity and suitability of analogues. In this case study, the *in silico* alerts relating to the reported weak *in vivo* endocrine activities were evaluated. Based on the working hypothesis that all analogues are converted to the same metabolite: pHBA and it's *in silico* alerts were also investigated. Firstly, the profilers that the OECD QSAR Toolbox highlights as pertinent for reproductive toxicity – i.e. the DART scheme, Oestrogen Receptor Binding, Retinoic Acid Receptor Binding – and the rtER Expert System were evaluated to examine the similarity among the category members (OECD, 2018). The *in silico* profiling results of the four parabens in the category and their common ester hydrolysis metabolite pHBA are listed in Table 5. MP, EP, PP, BP and the metabolite pHBA exhibited binding propensities for the oestrogen receptor; however, they were outside the applicability domain of the RAR-profiler. The ER profilers indicate that the short linear chain n-alkyl parabens displayed a small increasing trend in the order MP < EP < PP < BP regarding strength of binding affinity to the oestrogen receptor, as a function of alkyl chain length. The common metabolite pHBA was an outlier with respect to the parabens. These ER profilers only provide theoretical binding alert predictions, but do not translate into *in vivo* effects due to the absence of relevant exposure of the respective target organs. However, these predictions may support the category grouping.

In a second step, 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 (Kolsek et al., 2014). Therefore, the 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 non-binders. The docking simulations were used to characterize the binding propensities of short linear chain n-alkyl 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 (Kolsek et al., 2014). These thresholds were established according to a conservative approach as Kolšek and co-authors decided that the true-positive rate was 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 as well as their shared main metabolite pHBA are associated with a low binding probability class (green colour) for all receptors except a low intermediate outcome for the androgen receptor (AR) in antagonistic conformation (AR an.) (yellow colour). To provide comparison, five phytoestrogens Zeralenone (ZL, two stereoisomers), Coumestrol (CE), Genistein GE), Daidzein (DD), Apigenin (AG) were also analysed whose experimental characterisation highlighted affinities for the ER (Kuiper et al., 1998). All these chemicals are associated with docking scores highlighting an enhanced affinity (i.e. a lower docking scores) for the ERs and other targets (Table 6). We also added BPA that, as highlighted by the docking scores (Table 6) is characterized by stronger interactions with the estrogen receptors and a pronounced affinity for ERb. According to these comparison with control chemicals, the docking results suggest an overall negligible disrupting potential of short-chain parabens.

Overall, this *in silico* data support the lack of a relevant endocrine-related activity and the comparability of the data of the four parabens and the shared metabolite strengthens the selection of these category members.

The absence of relevant endocrine-related *in silico* activity is corroborated by *in vitro* data on parabens using oestrogen receptors. Routledge et al. (1998) tested MP, EP, PP, BP and 4-*n*-dodecyl paraben, as well as pHBA, in the *in vitro* recombinant yeast oestrogen screen and found the butyl, > propyl, > ethyl and > methyl ester to be weakly positive, whereas 4-n-dodecyl paraben and 4-hydroxybenzoic acid were without activity. Overall, the potencies of the parabens were several magnitudes below the endogenous natural substrate 17β-oestradiol, e.g. approximately 2,500,000-fold below for MP.

Okubo et al. (2001) found that the overall weak *in vitro* ooestrogenic activity of parabens increased in the order MP > EP > PP > BP > isopropyl paraben > isobutyl paraben by assaying oestrogen receptor dependent proliferation of human MCF7 breast cancer cells. Overall, endocrine-related *in vitro* activity was several magnitudes (10^{-5} to 10^{-7}

Table 5
In Silico Profilers Relevant to Reproductive Toxicity. Profiling results obtained from OECD QSAR Toolbox v 4.2

Chemical	DART scheme	Oestrogen 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

Table 6
Docking scores towards sixteen structures belonging to twelve nuclear receptors for pHBA and short chain parabens. 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. Zeralenone (ZL, two stereoisomers), Coumestrol (CE), Genistein GE), Daidzein (DD), Apigenin (AG), bisphenol-A (BPA).

	pHBA	MP	EP	PP	BP	ZL(S, E)	ZL(R, E)	CE	GE	DD	AG	BPA
AR	-6.0	-6.1	-6.3	-6.6	-6.8	-3.8	-5.9	-9.9	-9.1	-9.3	-8.9	-8.5
AR an.	-5.9	-5.9	-6.0	-6.3	-6.3	-7.4	-8.1	-9.6	-9.1	-9.0	-9.1	-8.6
ER α	-5.6	-5.7	-6.0	-6.5	-6.7	-9.8	-9.0	-9.4	-9.2	-9.4	-8.8	-8.2
ER α an.	-5.6	-5.7	-6.0	-6.4	-6.5	-7.9	-7.5	-9.1	-9.2	-9.9	-9.2	-8.5
ER β	-5.7	-5.8	-6.1	-6.4	-6.5	-10.5	-9.7	-9.8	-8.6	-8.7	-8.2	-8.2
ER β an.	-5.6	-5.7	-6.1	-6.4	-6.5	-9.5	-8.0	-9.6	-8.6	-8.8	-8.9	-8.2
GR	-5.7	-5.9	-6.1	-6.3	-6.4	-8.4	-8.2	-9.4	-9.0	-9.0	-8.8	-7.8
GR an.	-5.2	-5.4	-5.6	-5.8	-5.8	-8.2	-8.5	-8.0	-7.6	-7.3	-7.8	-7.4
LXR α	-5.5	-5.6	-5.9	-6.3	-6.4	-7.5	-8.3	-9.4	-8.7	-9.0	-8.8	-8.6
LXR β	-6.0	-6.1	-6.3	-6.7	-6.9	-6.9	-7.8	-9.7	-9.6	-9.9	-9.5	-8.0
PPAR α	-5.5	-5.5	-5.8	-6.4	-6.5	-6.6	-7.9	-8.3	-8.1	-7.7	-9.3	-7.9
PPAR β	-5.8	-5.7	-6.0	-6.0	-6.0	-7.0	-6.3	-9.3	-8.6	-7.9	-8.3	-7.8
PPAR γ	-5.3	-5.4	-5.9	-6.5	-6.7	-8.5	-8.0	-9.0	-8.1	-8.6	-9.3	-7.1
RXR α	-6.3	-6.0	-6.5	-6.9	-7.0	-7.6	-7.6	-8.7	-8.9	-9.5	-9.7	-7.9
TR α	-5.9	-5.9	-6.3	-6.7	-6.8	0.4	-2.2	-8.0	-9.6	-9.4	-9.6	-8.9
TR β	-5.7	-5.8	-6.1	-6.6	-6.7	-3.2	-3.5	-8.4	-9.6	-9.3	-9.4	-8.6

times) lower for all four n-alkyl parabens compared to that of natural endogenous substrates such as 17β-oestradiol.

c) Summary of the outcome of Tier 0 & determining the similarity hypothesis

The following conclusions can be made at the end of Tier 0:

TTC: The TTC concept could be applied for PP dermal exposure as the calculated exposure from the Creme Care and Exposure model scenario (d) leads to an aggregate exposure value of 0.014 mg/kg/day which is lower than the respective regulatory Cramer Class I TTC threshold value 0.03 mg/kg/day (EFSA 2019) or 0.042 mg/kg/day (Yang et al., 2017). However, to illustrate the process we continue with RAX.

Chemical structure similarity: In terms of Tanimoto similarity, MP (0.81), EP (0.93) and BP (0.94) were identified as the closest analogues to the target PP, which sits in between EP and BP in the homologous series.

Physicochemical properties: While there are some slight differences that may influence oral and dermal bioavailability with increasing side chain length, comparison of the physicochemical properties across the four parabens overall substantiates the suitability of the category as a similar set of homologues.

In vivo data: Overall, *in vivo* reproductive and developmental toxicity studies generally demonstrate no relevant adverse effects for MP, EP and BP up to 1000 mg/kg/day (Table 3). A POD of 1000 mg/kg/day is used in regulatory risk assessments for MP and EP (SCCP 2006; EFSA 2004). For BP, the SCCS selected a subcutaneous dose of 2 mg/kg/day as the POD for their safety evaluation. However, it is considered extremely conservative to define this dose as a legitimate NOEL/NOAEL for BP as it was the only dose tested in the respective rat study; a NOEL could well have been much higher than this in this study. Other state-of-the-art reproductive/developmental toxicity studies show a NOAEL up to and including 1000 mg/kg/day (Daston, 2004; Hoberman et al., 2008; Hubbard et al. 2020 (US NTP studies performed in 2011)). However, for the illustrative purposes of working through the complete 10-step framework, we have continued in this RAX based NGRA, with the knowingly conservative POD of 2 mg/kg/day for BP.

In vivo screening studies such as the uterotrophic assay indicated no or at most weak oestrogenic activity of the parabens, which presents no evidence of an endocrine-related adverse effect according to the OECD Framework for Testing and Assessment of Endocrine Disrupters (OECD, 2012).

Overall, the *in vivo* data support the suitability of the category as a similar set of homologues, with general low or no toxicity (Category III in terms of the scientific basis of the RAX (see Alexander-White et al., 2022)).

In silico alerts and docking simulations: Overall, *in silico* profiling and docking simulations indicate a homogenous profile of very weak binding activity for the receptors considered by the Endocrine Disruptome tool, further substantiating the suitability of the four parabens to form one category of similar analogues.

4. Tier 0 exit: step 4 → step 8 selection of a systemic toxicity point of departure

At this point, from the perspective of deriving a human relevant health guidance value, one could in principle move to Step 8, if one were confident that a POD could be selected using RAX that was representative and suitably conservative. Using the external exposure dose metric generated in Step 1, a risk assessment (Step 9) could be performed.

Step 8: Performing a RAX to derive a point of departure.

In the absence of data for PP, the assumption is made based on the chemical and biological similarity profiles in Tier 0, that the experimental POD for a paraben analogue can also be used conservatively for the PP in a MOS calculation. The question remains here as to whether PP is closer in biological similarity to MP and EP, each with an experimental

POD of 1000 mg/kg/day or whether PP is more similar to BP that has been assigned a considerably lower and highly unrealistic POD (2 mg/kg/day) from a Klimisch 3 rated study (SCCS, 2013). At this point, in this approach the most conservative of these values should be taken forward unless there is improved evidence of greater similarity to MP or EP to use a higher value.

Step 9 Performing a margin of safety evaluation

Using the POD for BP chosen by the SCCS (i.e. 2 mg/kg/day)) and dividing it by the worst case initial deterministic aggregate exposure estimate for PP in all cosmetics products (0.53 mg/kg bw/day), then the Margin of Safety (MOS) = 2/0.53 = 4. This MOS is clearly not sufficiently high to provide assurance at this point in the process as an acceptable MOS in this situation is typically expected to be 100 or more (WHO, 2005; SCCS, 2018).

However, this does not mean that parabens are unsafe, but at this point there is not enough data, evidence and realistic accuracy in the risk assessment, as we have taken highly worst case assumptions in both POD and exposure estimates. Therefore, further refinement is needed. When using the outcomes for PP exposure estimation using probabilistic modelling from the Creme Care and Exposure model, in the scenarios outlined in Step 1 above, the MOS are: a) 2/0.154 = 13; b) 2/0.053 = 38; c) 2/0.07 = 29, d) 2/0.014 = 143. However, it should be noted that when using the most realistic probabilistic exposure scenario (d), the MOS is acceptably high at 143.

Based on the deterministic exposure value, the margin of exposure is not sufficient to sustain the use scenario. A next step in the 10-step RAX framework is to move to Tier 1 (Steps 5 and 6) to define systemic bioavailability and assess whether there are areas of greatest mechanistic similarity between the target and the analogues to refine the NGRA further.

5. Tier 1

Step 5. Systemic bioavailability/ADME of PP and analogues

Understanding absorption, distribution, metabolism, and excretion (ADME) properties and the relative rate and extent of biotransformation across the short chain linear parabens is an important aspect in the examination of potential potency differences between analogues. 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 that is 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; Janjua et al., 2008; Ozaki et al., 2013; Moos et al., 2016). It is understood that after oral and dermal exposure, 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 the context of this case study. It is useful to consider in what amount PP penetrates across the skin following dermal application and absorbs into the systemic circulation and in what form.

a) *Ex vivo* absorption and metabolism in human skin

To best understand dermal bioavailability, potential first-pass metabolism in the skin should be considered as well as the dermal penetration (Manwaring et al., 2015). Experiments (following OECD guideline 428 for skin penetration (OECD 2004)) 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 (Table 7; Géniès et al., 2019). The mass balance recovery of the applied dose in the experiment with human skin was 91.8 ± 6.6% and the % of total radioactivity in the culture medium after 24 h was 66.0 ± 9.2%; 27.6% total radioactivity was found in the skin. The majority of radioactivity in the medium (42%) was pHBA, indicating the action of carboxylesterases in the skin and 5.6% was sulphonated PP. Twelve metabolites were detected (see Table 7 in Géniès et al., 2019) and only 0.2% of the administered dermal dose was measured after 24 h in receptor fluid as parent PP.

Table 7
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.

Compound	PHH μL/min/million cells	Liver S9 μL/min/mg protein	EpiSkin S9 μL/min/mg protein
MP	73.8 ± 8.0	129.1 ± 3.5	73.8 ± 8.0
EP	60.0 ± 3.4	94.2 ± 2.4	60.0 ± 3.4
PP	73.6 ± 15.5	84.4 ± 3.0	73.6 ± 15.5
BP	42.6 ± 0.2	105.0 ± 2.3	42.6 ± 0.2
pHBA	<1 ($t_{1/2}$ >180 min)	Not metabolised	Not metabolised
Naloxone (high)	9.8 ± 1.9	Not done	Not done
Midazolam (medium)	4.1 ± 0.5	Not done	Not done
Tolbutamide (low)	1.0 ± 0.1	Not done	Not done

These *ex vivo* and *in vitro* studies indicate that (i) very low amounts of parent paraben enter the systemic circulation after topical application of PP due to first-pass metabolism in the skin, and (ii) the major metabolite entering the systemic circulation is pHBA (Géniès et al., 2019). A pragmatic value of 1% can be used in risk assessment for the systemic delivery of parent PP by the dermal route. The measured value for skin penetration through human skin was 0.2 ± 0.2% but to allow for uncertainty and variability, given small sample numbers in this study and the potential for a very small amount residual in the skin tissue that could go on to be absorbed, a value of 1% is considered appropriately conservative.

b) Metabolism in *in vitro* liver S9, primary human hepatocytes and Episkin S9

Once entering the systemic circulation, a compound can be further metabolised by the liver. The *in vitro* intrinsic clearance was examined for the parabens in both Episkin (pool of 5 donors) and human liver S9 (pool of 200 donors, mixed gender), as well cryopreserved primary human hepatocytes (pool of 5 donors, mixed gender). Relative to reference compounds the intrinsic clearance (CL_{int}, *in vitro*) in primary human hepatocytes (PHH) indicated that all four parabens are high hepatic clearance compounds (Table 7). Assuming 1.29 million PHH is equivalent to 1 mg S9 (based on Lipscomb et al., 1998), 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.

As with metabolism in *ex vivo* skin (Géniès et al., 2019), in incubations with PHH, pHBA accounted for the majority of metabolite formed (Table 7), as depletion of parent was concomitant with an increase in pHBA formation. These findings are in accordance with others who have studied the metabolism of several parabens in human liver microsomes (Abbas et al., 2010).

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 metabolising enzymes that were responsible for paraben metabolism in liver- and skin-based models (Eilstein et al., 2019). The rate of metabolism of the four parabens was much higher (between 70- and 210-fold

higher) in liver than in EpiSkin S9 (Table 7). 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 that the metabolism of the parabens in EpiSkin S9 is mediated by CES2. There were two detected metabolites common to PHH, liver S9 and EpiSkin S9, namely pHBA and a direct sulfate conjugate of the parabens S9. 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 is expected considering the much lower abundance and activities of CYP enzymes in skin compared to the liver (Hewitt et al., 2013). When pHBA was incubated with liver or EpiSkin S9, it was not metabolised and no conjugates were detected. This finding again indicates that pHBA is the major metabolite via the action of esterases and cytochrome P450 enzymes are insignificant for parabens metabolism.

c) Metabolism in plasma

In addition to undergoing metabolism in the liver, some esters are hydrolysed by esterases in the plasma (Fu et al., 2016). As 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. The degree of plasma protein binding was high and increased with increasing paraben chain length (Table 8). The fraction bound suggests that the free fraction *in vivo* could vary between 26% for MP to only 4% for BP. Despite the high extent of protein binding, this did not prevent the parabens from being metabolised, suggesting that the binding affinity was low enough to release the compound for metabolism. The observed binding of p-HBA to plasma proteins was much lower than that of any of the parent chemicals.

Step 6. Supporting a Similar Mode/Mechanism of Action (MOA) hypothesis

The working hypothesis for this category is that, based on their highly similar chemical structure, the target chemical PP will have similar biological activity and bioavailability to the source chemicals MP, EP, and BP.

Table 8
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	Stability control			
	PPB [%]	% Recovery	% parent remaining after 1 h	nM 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

The key aspects of the hypothesis are as follows:

i) Similar chemical structure and physicochemical characteristics will result in similar bioavailability, metabolism, and reactivity, which results in similar biological and functional effects.
ii) The available *in vivo* systemic toxicity data generally demonstrate similar biological activity across the category.
iii) 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 and major primary metabolite, pHBA, and similar corresponding short linear chain alcohols. At the levels of exposure to parabens in cosmetics, the alcohols generated are not of concern toxicologically.
iv) The rate and extent of ester hydrolysis is similar across parabens, resulting in similar exposures to the common metabolite pHBA, which is not toxic
v) Chain length differences across the parabens may result in a predictable potency trend in observed effects across category members with increasing alkyl chain length e.g. in *in vitro* assays and uterotrophic assays etc.

To investigate and support the hypothesis further and to explore biological similarity, US EPA ToxCast data were analysed.

5.1. Bioactivity in ToxCast (potential mode of action (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 (US 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 MoA and potential similarities.

Initially, a structure similarity search utilising Accelrys Isentris (v4.0) 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 pulled back from GRASP (Graphical Structure Project), a proprietary P&G platform, 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 in Step 1 by Tanimoto and expert chemical review) and the common metabolite pHBA.

Analysis of the ToxCast data associated with MP, EP, PP, BP and pHBA was undertaken. The analysis focused on assay hits with no flags as reported by the US EPA. Flags associated with response data from ToxCast assays indicates potential issues with the fit model (false positive/false negative) as identified by the US 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 9.

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.

Due to this convergence, the ToxCast oestrogen receptor model was explored further (personal communication with US EPA). Full details of the oestrogen receptor model are described elsewhere (Browne et al.,

Table 9
ToxCast assay hit counts for parabens.

Name	CAS	ToxCast Chemical ID	Similarity Cutoff (Isentris)	Result Count all assays	Result Count hits	% 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

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 Fig. 3) demonstrate that the rank order of potency, albeit low, for oestrogen receptor activity is MP < EP < PP < BP, with the reference substance showing, as expected, much greater oestrogen receptor activity overall.

	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 AC_{10} and AC_{50} values listed in Fig. 3 were derived by R. Judson at the US EPA (personal communication). BP is associated with the lowest concentrations for both AC_{10} and AC_{50} in comparison to the other parabens. Thus, it is assigned a potency of 1 relative to the other category members (see Table 10).

ToxCast data traditionally rely on the concentration of chemical associated with 50% of maximum activity, i.e. AC50. However, because this assay response could reflect agonist effects on the oestrogen receptor, increasing concentrations could trigger increasing oestrogenic 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.

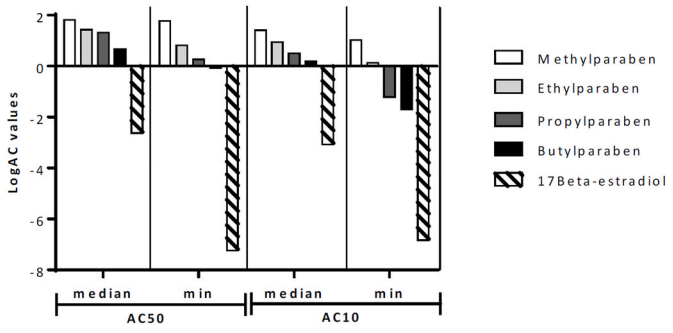


Fig. 3. Relative oestrogen receptor bioactivity in ToxCast. AC values are shown in the table below the graph.

Table 10
Calculation of potency scaling factors from ToxCast oestrogen receptor activity data AC10.median. Calculated Scaling (potency) Factor*.

	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

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

5.2. Toxicogenomics data

The results of the toxicogenomics analyses using MCF7 cells (Fig. 4), indicate that each of the parabens is able to elicit changes in the expression of a large number of genes (FDR <0.05, fold change ± 1.2>), as compared to controls, particularly at the highest dose tested. 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 because these cells express multiple nuclear hormone receptors as well as other regulatory proteins. 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) (FDR <0.05, fold change ± 1.2>). This is shown in the Eisen diagram heat map (Fig. 4) 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. Fig. 4 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 four 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). 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. 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 (details on these results are published in the OECD IATA report: ENV/JM/MONO(2020)16 OECD Series on Testing and Assessment No. 320).

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

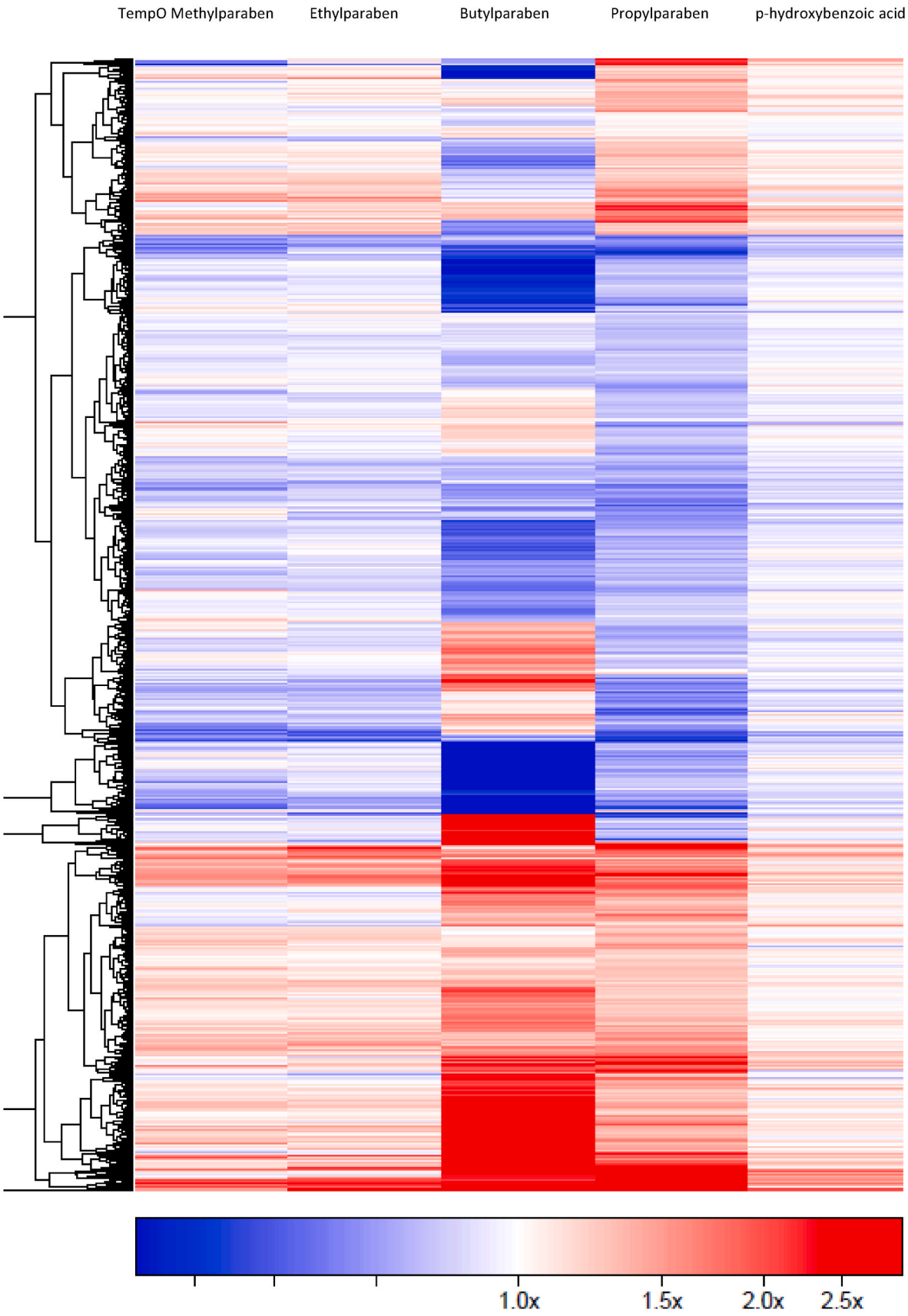


Fig. 4. Heat map of the genes whose expression was modified in MCF7 cells treated with MP, EP, PP and BP. Up-regulated genes in red; down-regulated genes in blue.

pathway enrichment. Looking broadly across the four parabens, there was 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 and MTORC1 signalling.

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 *in vitro*, and more specifically provide evidence that the parabens share potential MoAs.

6. Tier 1 exit: step 6 → step 8 selection of a systemic toxicity point of departure

At the end of Tier 1, ADME data indicate similarity of bioavailability for the parabens and data from ToxCast and toxicogenomics data in MCF7 cells above further increases confidence in the biological similarity of the analogues in *in vitro* assays and increases the confidence in the assumption at the end of Tier 0, about the use of the highly conservative POD for PP as read across from the POD currently in use for BP (2 mg/kg/day). The data suggest that PP is likely to be less biologically active than BP and the relative potency factors from the oestrogen receptor assays can be used in the final risk assessment. Further evidence on parabens activity using targeted testing and refinement to exposure estimates can still be made if we continue to Tier 2, with the use of PBK modelling data to determine internal dose metrics.

7. Tier 2

In Tier 2, toxicogenomics analysis and data available in ToxCast suggest that further targeted testing could be useful in exploring relative potency and biological similarity further. The bioavailability data also suggest that a PBK model can be built using available data on MP, EP and BP, which can then be used to generate estimate of PP kinetics and internal dose metrics. Therefore, we can progress to using Steps 7a and 7b.

Step 7a Perform Targeted Testing: Exploring CALUX assays with parabens

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, the low binding alerts and binding activity of the parabens observed in the molecular docking and *in silico* profilers and the bioactivity (ToxCast) data already gathered, specific CALUX® transactivation assays (OECD 2016) were selected to examine the similarities and differences in the endocrine activity of parabens. As 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. As such, interaction with receptors for oestrogen-, androgen-, thyroid signalling and steroidogenesis (EATS) are relevant to potential MoAs based on *in vitro* 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. The outcomes of the CALUX assays are listed below.

a) Cytotoxicity assay

In the cytotoxicity CALUX assay (data not shown), 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 1×10^{-3} M.

b) Oestrogen and Androgen receptor assays

All parabens showed oestrogenic activity in the absence of rat liver S9, but no anti-oestrogenic activity was observed (Fig. 5). 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. These results were in agreement with observations by Watanabe et al. (2013) on 17 parabens with the ERα and ERβ receptors.

The AR CALUX assay showed that none of the compounds had androgenic activity, while they did show anti-androgenic activity (Fig. 6). 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.

c) Thyroidogenic activity

No significant thyroidogenic activity was detected for any of the compounds, and anti-thyroidogenic activity was observed for MP. Also, for the second thyroid-related assay, hTPO inhibition, little activity was observed. 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.

d) Steroidogenic activity

All four parabens affected steroidogenesis following exposure of H295R cells and subsequent quantification of 17beta-estradiol and/or testosterone production using the ERα and AR CALUX bioassay (OECD 2011). 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'.

7.1. Summary of the EATS assays

Importantly, incubations with S9 in all cases decreased bioactivity in the EATS panel (Table 11). 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 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

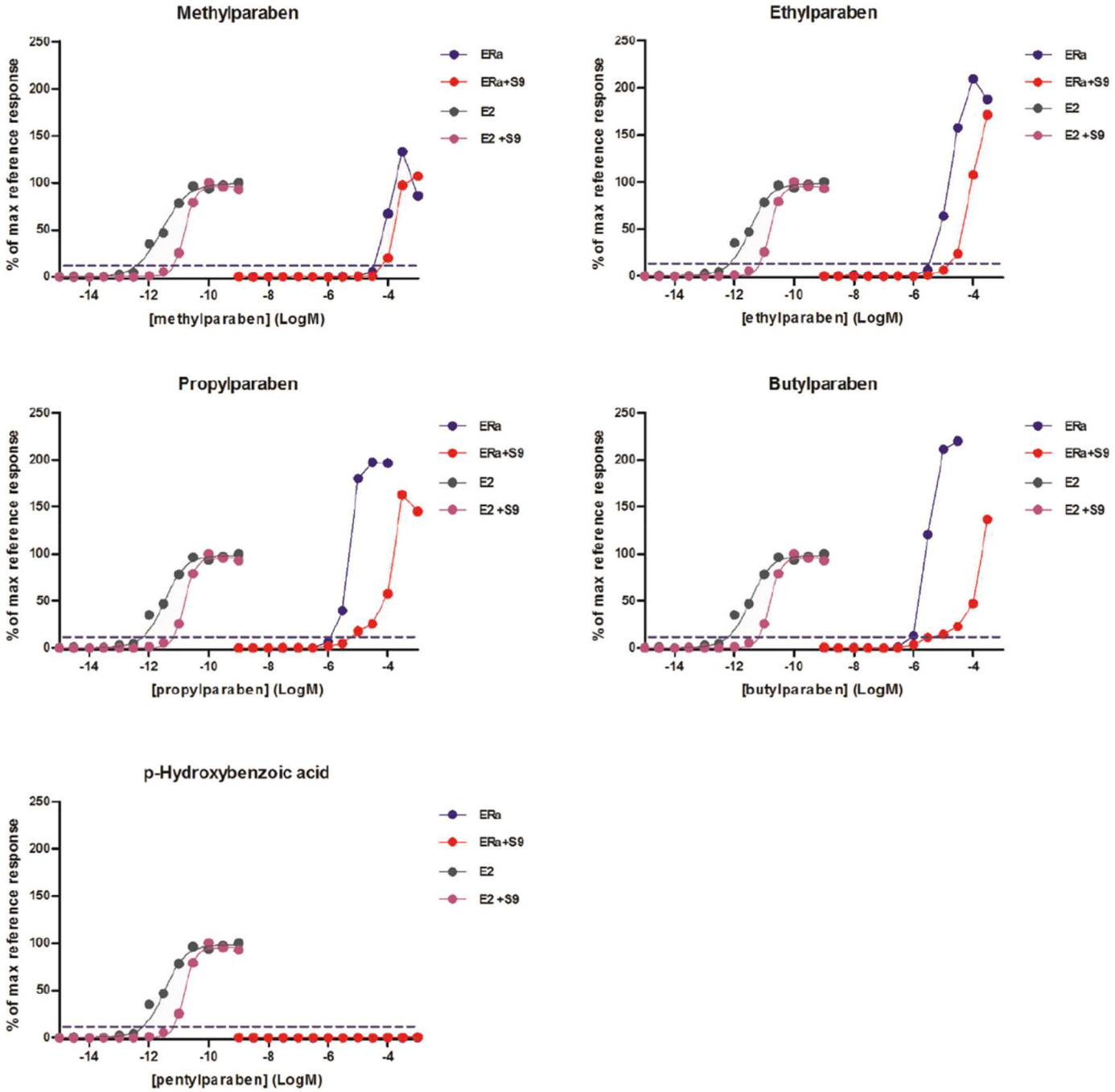


Fig. 5. ERα CALUX results. Receptor activation (% of maximum) is plotted against compound concentration (LogM) final in well. The assay was performed in the absence (blue) and presence of metabolic enzymes (rat liver S9 fraction). Samples were prepared in triplicate and cells were exposed to test substance for 24h. 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).

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 (Fig. 5), while being antagonists in the AR assay at high concentrations (Fig. 6). This linked activity has been noted before in other endocrine active substances. 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 *in vitro* with increasing chain length, suggesting a trend in potency across the category. The results from the EATS assays, with and without metabolic activity, supported the earlier findings of ER activity from *in silico* alerts and ToxCast data, where for the latter differences are greater in the absence of metabolism. However, it has to be emphasised that in all EATS assays, parabens are many orders of magnitude less potent compared to the natural oestrogen 17β-estradiol (Golden et al., 2005).

Physiologically-based kinetic (PBK) models are mathematical

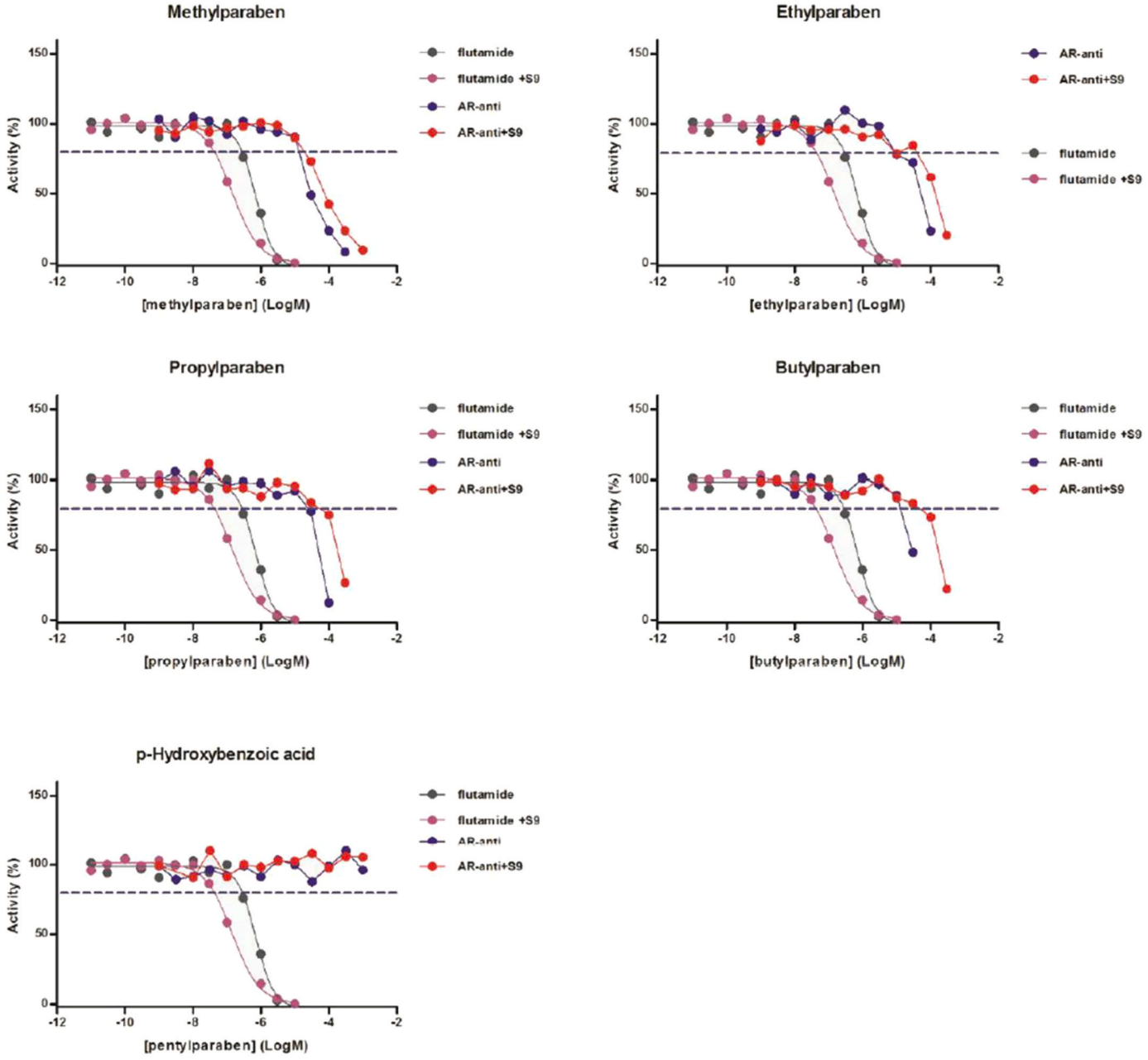


Fig. 6. Anti-AR CALUX results. Activity (% compared to EC50-agonist response) is plotted against compound concentration (LogM) final in well. The assay was performed in the absence (blue) and presence of metabolic enzymes (rat liver S9 fraction). Samples were prepared in triplicate and cells were exposed to test substance for 24h. 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).

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. 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 PBK models provides a particularly useful framework for conducting cross species extrapolations. The application of PBK 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. 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. To apply a PBK 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 PBK model are then

Table 11
Summary of EATS testing results. PC10 (for agonistic tests)/PC20 (for antagonistic tests) values are shown in -Log M; the color indicates the potency (yellow < orange < red).

Step 7b. *Biokinetic refinement*

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-) estrogenic and (anti-) androgenic assays										
ER α CALUX	-4.5	-4.2	-5.5	-4.8	-6.0	-5.1	-6.0	-5.0	>	>
anti-ER α CALUX	>	>	>	>	>	>	>	>	>	>
AR CALUX	>	>	>	>	>	>	>	>	>	>
anti-AR CALUX	-4.9	-4.7	-4.7	-4.4	-4.5	-4.2	-4.9	-4.3	>	>
Thyroidogenic assays										
TR β CALUX	>	>	>	>	>	>	>	>	>	>
anti-TR β 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

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.

The details of the PBK 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 the OECD IATA report for propylparaben case study (ENV/JM/MONO(2020)16 OECD Series on Testing and Assessment No. 320). An overview of the model structure is shown in Fig. 7.

Various guidance documents for the application, use, best practice

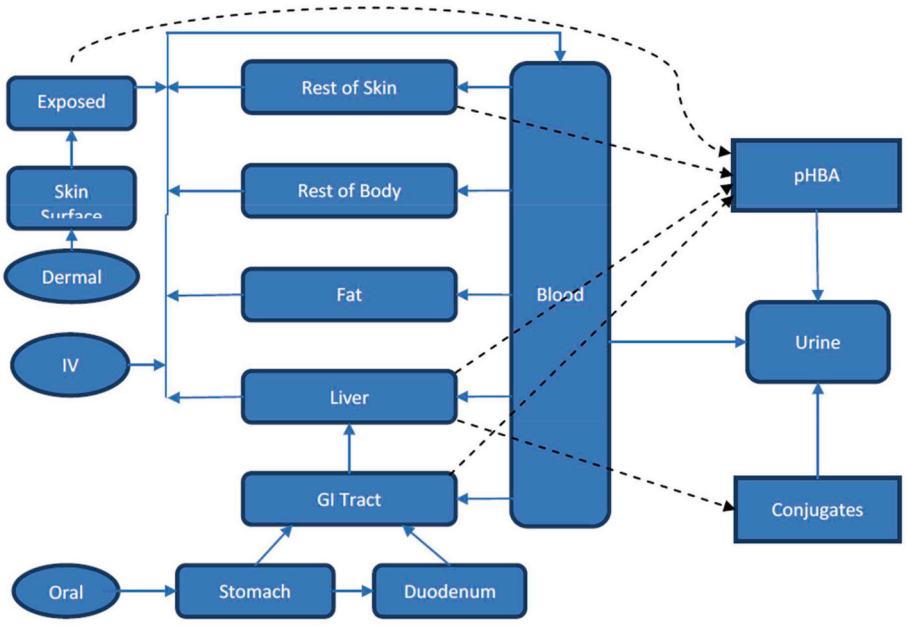


Fig. 7. PBK model schematic for parabens. Parent compound may be hydrolysed in the liver, skin, and gastrointestinal (GI) tissue, and conjugated (glucuronidation and sulfation) in the liver. Parent and metabolites may be excreted in urine. A fat compartment is included as a storage tissue.

and reporting of PBK models have been published (WHO, 2010; U.S. EPA, 2006; U.S. FDA, 2018; European Medicines Agency, 2016). Additionally, in order to address the credibility of PBK models for new chemicals on the market for which *in vivo* data cannot be generated for evaluation, an international effort at the OECD has delivered a guidance document the characterisation, validation and reporting of physiologically based kinetic models (PBK) for regulatory purposes (Sachana, 2019; OECD, 2021). A number of recent reviews of PBK modelling in environmental risk assessment are available (Clewell 2005; Clewell and Clewell 2008; Campbell et al., 2012; Clewell et al., 2014) and a paper on parabens PBK modelling (Campbell et al., 2015).

a) PBK modelling in consumer exposure to parabens from dermally applied cosmetics

Exposure estimates generated using the Creme Global exposure model were used as input to the PBK model (Table 12). For PP, the internal exposure estimates were Cmax of 0.022 μM, the AUC was 0.370 μmole*h/L and the Cavg was 0.016 μM from the SCCS deterministic consumer exposure estimates; Cmax of 0.018 μM, the AUC was 0.310 mmol*h/L and the Cavg was 0.013 μM from the Crème deterministic (worst case) consumer exposure estimates; and Cmax of 0.0006 μM, the AUC was 0.010 μmole*h/L and the Cavg was 0.0004 μM from the Crème probabilistic (realistic) consumer exposure estimates.

b) PBK modelling in rats after subcutaneous exposure to parabens

Based on read-across from BP, the conservative POD of 2.0 mg/kg/day is used for risk assessment (although much higher estimates exist) 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 is shown in Fig. 8. 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 values representing the POD are: Cmax 2.1 μM, AUC 3.0 μmole*h/L and Cavg 0.13 μM.

Step 8 Performing a RAX to derive a point of departure (POD)

At the end of Tier 2, there was no further strong evidence at this time that PP was toxicologically more similar to MP (with a POD of 1000 mg/kg/day) than to BP (with a conservative POD of 2 mg/kg/day selected by the SCCS). Given data available post the 2013 animal testing ban was not used in principle in this NGRA, it was concluded that the more conservative POD of 2 mg/kg/day for BP would have to be used as a comparative POD for PP in this RAX-based risk assessment. The benefit of using this lower value is however, that due to this highly conservative choice there is high confidence that the overall outcome is protective of human safety. It is assumed that in reality the POD is much higher than 2 mg/kg/day as explained earlier.

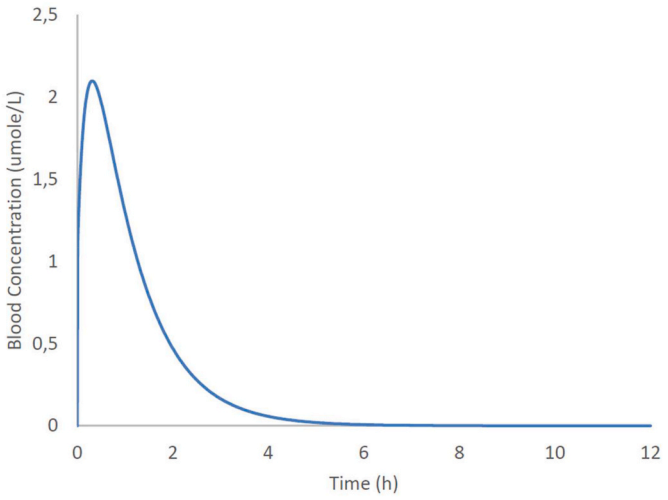


Fig. 8. Rat plasma time-course simulation of exposure in the study. Rats were injected subcutaneously with 2 mg/kg/day BP. Only one day is shown as the clearance of parent compound is complete in less than 12 h..

Step 9 Next Generation Risk Assessment: Perform a Margin of Internal Exposure (MoIE) assessment using PBK data

From the PBK modelling in Step 7b, it has been concluded that external SC exposure in rats to 2 mg/kg/day of BP (the POD from step 8 determined after Tier 0) results in an internal exposure C_{max} of 2.1 μM. Similarly, using PBK modelling, the human exposure simulation suggests an internal exposure of 0.022 μM to the target chemical PP (Table 12) when using deterministic values. When using the refined probabilistic consumer exposure evaluation for the realistic exposure scenario (i.e. scenario d), the human exposure simulation suggests an internal exposure of 0.018 μM for conservative exposure assumptions (scenario a) and 0.0006 μM for realistic exposure assumptions (scenario

Table 12
Summary of human plasma data for the PBK simulations of exposures estimated with the Creme Care and Exposure modelling tool.

Chemical	Exposure (Creme Global)			Cmax	AUC	Cavg
	Scenario	mg/kg/d	μg/cm2	μmole/L	μmole*h/L	μmole/L
MP	a	0.368	0.80	1.4E-02	2.8E-01	1.2E-02
EP	a	0.262	0.57	1.1E-02	1.8E-01	7.7E-03
PP	a	0.154	0.33	6.4E-03	1.1E-01	4.6E-03
BP	a	0.091	0.20	3.4E-03	6.1E-02	2.5E-03
MP	b	0.111	0.24	4.1E-03	8.4E-02	3.5E-03
EP	b	0.059	0.13	2.6E-03	4.2E-02	1.7E-03
PP	b	0.053	0.11	2.2E-03	3.8E-02	1.6E-03
BP	b	0.037	0.08	1.4E-03	2.5E-02	1.0E-03
MP	c	0.183	0.40	6.8E-03	1.4E-01	5.8E-03
EP	c	0.078	0.17	3.4E-03	5.5E-02	2.3E-03
PP	c	0.07	0.15	2.9E-03	5.0E-02	2.1E-03
BP	c	0.045	0.10	1.7E-03	3.0E-02	1.3E-03
MP	d	0.059	0.13	2.2E-03	4.5E-02	1.9E-03
EP	d	0.019	0.04	8.0E-04	1.3E-02	6.0E-04
PP	d	0.014	0.03	6.0E-04	1.0E-02	4.0E-04
BP	d	0.018	0.04	7.0E-04	1.2E-02	5.0E-04

Table 13
Margin of Internal Exposures (MoIE) using PBK modelling outputs for the POD and estimated human exposures of parabens in cosmetic products.

Following a deterministic consumer exposure estimate, the internal MoIE is calculated:			
POD	Internal exposure	Relative Potency	MoIE
Cmaxrat for BP: 2.1 μM	Cmaxhuman PP:2.2 × 10^{-2} μM	Factor for PP:0.37	MoIE = 2.1/(2.2 × 10-2* 0.37) = 258
Following probabilistic consumer exposure estimates for worst case and realistic scenarios, the internal MoIE is calculated as: Creme model, Tier 1 deterministic (worst case scenario a)			
POD	Internal exposure	Relative Potency	MoIE
Cmax rat for BP: 2.1 μM	Cmax human PP: 1.8E-2 μM	Factor for PP: 0.37	MoIE = 2.1/(1.8E-2 * 0.37) = 315
Creme model, Tier 2 probabilistic (realistic scenario d)			
POD	Internal exposure	Relative Potency	MoIE
Cmax rat for BP: 2.1 μM	Cmax human PP: 6.0E-4 μM	Factor for PP: 0.37	MoIE = 2.1/(6.0E-4 * 0.37) = 9459

d) (see Table 12).

Based on the relative potency information on the parabens that was gained in the NAM evaluations in Step 5, the internal exposure can further be adjusted for relative 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 broadly similar but activity increases with increasing alkyl chain length, and quite markedly from propyl to butyl. This was particularly demonstrated based on NAM evaluations of the weak endocrine activity of parabens, in particular in ER activity evaluated in ToxCast assays. As the risk assessment endpoint is reproductive/developmental toxicity, endocrine activity may be a potential MoA. Therefore, the relative ER bioactivity based on ToxCast AC10 values (see Table 10) 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 MoIE is calculated

Table 14
Assessing the level of confidence for the NAMs used in the parabens case study.

Data type/Endpoint	Assumptions	Level of confidence (low, medium, high)	Comments
In vivo data	The POD is appropriately conservative for the target substance	High	*In vivo* study was chosen because was used by SCCS. Study ranked Klimisch score 3 (non- guideline, no dose-response, single dose, no effects seen). The POD derived from a single dose SC study is 2 mg/kg/day for BP which is very conservative compared to other *in vivo* studies.
Exposure data	The exposure estimate finally used in the NGRA does overestimate consumer exposure in reality	High	Predicted exposure using a deterministic estimate that is highly conservative and much more than consumers are exposed to in reality
NAM			
Molecular Docking/ ER activity	These docking simulations can characterize the binding propensities of short linear chain parabens and their common ester hydrolysis metabolite pHBA towards twelve nuclear receptors	High	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	ToxCast can increase confidence in the similarities of the structurally related chemicals in the category and inform on MoA & potency	Medium	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/pH BA activity	pHBA, the main metabolite of parabens, does not contribute to the observed low reproductive toxicity potential associated with exposure to parabens	Medium	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	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.	High	The assay was perfomed 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	Toxicogenomic data can inform on the gene expression changes and support the identification of the specific biologic activity of parabens.	High	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 expression 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 in regard to biological coverage because only one cell line was used.
PBK	PBK 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.	Medium	A PBK 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.

Medium, high level of confidence = uncertainty results minor or major conservatism in the safety assessment (i.e. overestimation of risk).
Low level of confidence = uncertainty results in minor or major concerns in the safety assessment (i.e. underestimation of risk).
*Key to direction and magnitude.

using the equation:

MoIE = Cmaxrat BP/[(Cmaxhuman PPx (Relative Potency of PP/BP)]

The resulting MoIEs are shown in Table 13.

When using deterministic values, the resulting MoIE is 258, whereas when using the probabilistic Tier 1 and Tier 2 consumer exposure estimates according to the Creme Global model, the MoIEs are 315 and 9459, respectively.

A MoIE differs from a traditional margin of safety (MOS) in that it is calculated as the ratio of a measure of internal exposure, such as blood/plasma concentration or target-tissue dose, rather than a measure of external exposure concentration, total bolus dose or ingested dose (Bessems et al., 2017). Thereby, the uncertainty in the risk assessment is considerably reduced and the default uncertainty factor of 4 for interspecies differences in toxicokinetics can be replaced (WHO, 2010). Thus, a MOIE of 25 is considered equivalent to the default MOS of 100, but with greater precision for the target chemical. As all MoIEs derived in this case study were largely above 25, they were considered sufficiently protective.

Step 10 Assessing the Level of Confidence in the Risk Assessment

Overall, the level of confidence was considered high (Table 14) as the evidence provided by the ADME and the toxicodynamic properties points to low/no toxicity based on the considered exposure scenario.

8. Conclusion

This case study for the target chemical propylparaben demonstrates the practical application of the 10-step RAX framework for NGRA, as described in Alexander-White et al. (2022). This complements an accompanying case study for caffeine, which followed the same approach (Bury et al., 2020) and has been reviewed by the OECD (2020).

The data provided for parabens, illustrates how read-across can be used to fill the data gaps on reproductive/developmental toxicity as a suspected pivotal toxicity endpoint for the target chemical PP. Source chemicals MP, EP and BP were included in a category approach to evaluate chemical and biological similarity and explore relative potency trends across the category using *in vitro* assay data particularly related to oestrogenic activity, as suspected biological activity. Multiple data streams were integrated in an IATA (Integrated Approach to Testing and Assessment) 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 (see Gazin et al., 2013), the information gathered has shown that non-animal methods can be used today to support the safety of short linear n-alkyl chain parabens as used in cosmetic products, even when highly conservative assumptions are made in the safety assessment.

Overall, the parabens are substances of low toxicity and all of the good quality studies indicate 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 (ECHA REACH dossier; Gazin et al., 2013 reporting studies performed at Ricerca Biosciences). There was no evidence of oestrogenic activity at any *in vivo* dose, and no effects on reproductive organs or function, which fully supports the weak ER-agonist activity of PP determined in various *in vitro* systems (i.e. ER-binding assays, CALUX data, etc.). The experimental NOAEL for PP in repeat dose OECD guideline studies is 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 hydrolysed by esterases 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. Overall, there was no evidence of any adverse effects up to the limit dose of 1000 mg/kg/day (Gazin et al., 2013; studies performed in 2012).

Based on conflicting results from the literature, there remain concerns that the parabens possess oestrogenic activity *in vivo*. However, there is little convincing evidence of this and oestrogenic activity observed *in vitro* is extremely weak (several magnitudes lower at maximum concentrations compared to the endogenous substrate 17beta-estradiol). Sporadic reports of alleged *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, where metabolism is not at play, the toxicological relevance for human safety continues to be unlikely. To date there is no *in vivo* evidence of adverse effects in humans 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.

In conclusion, as demonstrated in this case study, NAM data can provide useful information to facilitate the selection of the most appropriate analogue from a homologous series of chemicals to read across to a target category member. In addition, NAMs can be used in principle 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. The margin of internal exposure derived here was shown to be protective of human health.

Funding body information

This work was funded through the Long Range Science Strategy (LRSS) programme of Cosmetics Europe (https://www.lrsscosmeticseurope.eu).

CRediT authorship contribution statement

Gladys Ouedraogo: Conceptualization, Writing – review & editing. **Camilla Alexander-White:** Writing – original draft. **Dagmar Bury:** Conceptualization, Writing – review & editing. **Harvey J. Clewell:** Writing – review & editing. **Mark Cronin:** Methodology, Writing – review & editing. **Tom Cull:** Writing – review & editing. **Matthew Dent:** Conceptualization. **Bertrand Desprez:** Conceptualization, Project administration. **Ann Detroyer:** Conceptualization, Writing – review & editing. **Corie Ellison:** Data curation, Formal analysis, Methodology, Writing – review & editing. **Stefania Giammanco:** Writing – review & editing. **Eric Hack:** Data curation, Formal analysis, Methodology, Writing – review & editing. **Nicola J. Hewitt:** Methodology, Data curation, Formal analysis, Writing – review & editing, Project administration. **Gerry Kenna:** Conceptualization, Writing – review & editing. **Martina Klaric:** Writing – review & editing, Project administration. **Reinhard Kreiling:** Writing – review & editing. **Cathy Lester:** Writing – review & editing, Data curation, Formal analysis. **Catherine Mahony:** Writing – review & editing. **Enrico Mombelli:** Data curation, Formal analysis, Writing – review & editing. **Jorge Naciff:** Methodology, Data curation, Formal analysis, Writing – review & editing. **John O'Brien:** Writing – review & editing. **Andreas Schepky:** Writing – review & editing. **Sarah Tozer:** Data curation, Formal analysis, Writing – review & editing. **Bart van der Burg:** Methodology, Data curation, Formal analysis, Writing – review & editing. **Barbara van Vugt-Lussenburg:** Methodology, Data curation, Formal analysis, Writing – review & editing. **Sharon Stuard:** Writing – review & editing. **Cosmetics Europe:** Resources, Project administration.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

The authors would like to thank Annette Bitsch (Fraunhofer), Joan Fisher (P&G), Bastian Selman (P&G) for their help in the preparation of data included in this manuscript.

References

Abbas, S., Greige-Gerges, H., Karam, N., Piet, M.H., Netter, P., Magdalou, J., 2010. Metabolism of parabens (4-hydroxybenzoic acid esters) by hepatic esterases and udp-glucuronosyltransferases in man. Drug Metabol. Pharmacokinet. 25 (6), 568–577.

Alexander-White, C., Mahony, C., Bury, D., Cronin, M., Dent, M., Hack, E., Hewitt, N., Kenna, G., Naciff, J., Ouedraogo, G., Shepky, A., 2022. A 10-step framework for use of read-across (RAX) in next generation risk assessment (NGRA) for cosmetics safety assessment. Manuscript submitted to. Jan 3 Regul. Toxicol. Pharmacol., 105094. https://doi.org/10.1016/j.yrtph.2021.105094. Epub ahead of print. PMID: 34990780.

Aubert, N., Ameller, T., Legrand, J.J., 2012. Systemic exposure to parabens: pharmacokinetics, tissue distribution, excretion balance and plasma metabolites of [^{14}C]-methyl-, propyl- and butylparaben in rats after oral, topical or subcutaneous administration. Food Chem. Toxicol. 50 (3–4), 445–454.

Berggren, E., White, A., Ouedraogo, G., Paini, A., Richarz, A.N., Bois, F.Y., Exner, T., et al., 2017. Ab initio chemical safety assessment: a workflow based on exposure considerations and non-animal methods. Comput Toxicol 4, 31–44. https://doi.org/10.1016/j.comtox.2017.10.001.

Bessems, J.G.M., Paini, A., Gajewska, M., Worth, A., 2017. The margin of internal exposure (MoIE) concept for dermal risk assessment based on oral toxicity data - a case study with caffeine. Toxicology 392, 119–129. https://doi.org/10.1016/j.tox.2017.03.012.

Browne, P., Judson, R.S., Casey, W.M., Kleinstreuer, N.C., Thomas, R.S., 2015. Screening chemicals for oestrogen receptor bioactivity using a computational model. Environ. Sci. Technol. 49 (14), 8804–8814. https://doi.org/10.1021/acs.est.5b02641.

Bury, D., Alexander-White, C., Clewell III, H.J., Cronin, M., Desprez, B., Detroyer, A., Efremenko, A., Firman, J., Hack, E., Hewitt, N.J., Kenna, G., Klaric, M., Lester, C., Mahony, C., Ouedraogo, G., Paini, A., Schepky, A., 2020. New Framework for a Non-animal Approach Adequately Assures the Safety of Cosmetic Ingredients – A Case Study on Caffeine. Manuscript submitted to *Regul Toxicol Pharmacol* Special Issue.

Campbell, J.L., Yoon, M., Clewell, H.J., 2015. A case study on quantitative in vitro to in vivo extrapolation for environmental esters: methyl-, propyl- and butylparaben. Toxicology 332, 67–76. https://doi.org/10.1016/j.tox.2015.03.010.

Campbell Jr., 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.N. (Eds.), Computational Toxicology. Volume 1, Methods Mol Biol., pp. 439–499.

Clewell, H.J., 2005. Use of mode of action in risk assessment: past, present, and future. Toxicol Pharmacol 42 (1), 3–14. https://doi.org/10.1016/j.yrtph.2005.01.008.

Clewell, H.J., Clewell, R.A., Andersen, M.A., 2014. Physiologically based pharmacokinetic modeling. In: Hayes, W.A. (Ed.), Principles and Methods of Toxicology, sixth ed. CRC Press, New York.

Clewell, R.A., Clewell, H.J., 2008. Development and specification of physiologically based pharmacokinetic models for use in risk assessment. Regul. Toxicol. Pharmacol. 50 (1), 129–143. https://doi.org/10.1016/j.yrtph.2007.10.012.

Cosmetics Ingredient Review (CIR), 2008. Final amended report on the safety assessment of methylparaben, ethylparaben, propylparaben, isopropylparaben, butylparaben, isobutylparaben, and benzylparaben as used in cosmetic products. Int. J. Toxicol. 27 (Suppl. 4), 1–82. https://doi.org/10.1080/10915810802548359.

Danish EPA, 2009. Dermal absorption of pesticides – evaluation of variability and prevention. Pesticides Research 124.

Daston, G.P., 2004. Developmental toxicity evaluation of butylparaben in Sprague-Dawley rats. Birth Defects Res. Part B Dev. Reproductive Toxicol. 71 (4), 296–302. https://doi.org/10.1002/bdrb.20016.

Dearden, J., Worth, A., 2007. In: Silico Prediction of Physicochemical Properties. EC JRC Scientific and Technical Reports. Eur 23051 En.

EACH dossier for Propyl Paraben Dossier. Available at: https://echa.europa.eu/fr/registration-dossier/-/registered-dossier/13890/7/9/1.

European Food Standards Authority (EFSA), 2004. Opinion of the Scientific Panel on Food Additives, Flavourings, Processing Aids and Materials in Contact with Food on a Request from the Commission related to para hydroxybenzoates (E 214-219). EFSA J. 83, 1–26.

EFSA, 2019. Guidance on the use of the threshold of toxicological concern approach in food safety assessment. EFSA J. https://doi.org/10.2903/j.efsa.2019.5708. Guidance adopted April. 2019.

Eilstein, J., Grégoire, S., Fabre, A., Poupon, A., Arbey, E., Géniès, C., Duplan, H., Rothe, H., Ellison, C., Cubberley, R., Schepky, A., Klaric, M., Hewitt, N.J., Jacques-Jamin, C., 2019. 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. J. Appl. Toxicol. 40 (3), 416–433. https://doi.org/10.1002/jat.3914.

Ellison, C.A., Blackburn, K.L., Carmichael, P.L., Clewell, H.J., Cronin, M.T.D., Desprez, B., Escher, S.E., 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. 103, 63–72. https://doi.org/10.1016/j.yrtph.2019.01.016.

European Medicines Agency (EMA), 2016. Guideline on the Qualification and Reporting of 5 Physiologically Based Pharmacokinetic (PBPK) Modelling 6 and Simulation. EMA/CHMP/458101/2016. https://www.ema.europa.eu/en/documents/scientific-guideline/draft-guideline-qualification-reporting-physiologically-based-pharmacokinetic-pbpk-modelling_en.pdf.

Fagerberg, L., Hallstrom, B.M., Oksvold, P., Kampf, C., Djureinovic, D., Odeberg, J., Habuka, M., et al., 2014. Analysis of the human tissue-specific expression by genome-wide integration of transcriptomics and antibody-based proteomics. Mol. Cell. Proteomics 13 (2), 397–406. https://doi.org/10.1074/mcp.M113.035600.

Fisher, J.S., Turner, K.J., Brown, D., Sharpe, R.M., 1999. Effect of neonatal exposure to oestrogenic compounds on development of the excurrent ducts of the rat testis through puberty to adulthood. Environ. Health Perspect. 107 (5), 397–405. https://doi.org/10.1289/ehp.99107397.

Fu, J., Pacyniak, E., Leed, M.G.D., Sadgrove, M.P., Marson, L., Jay, M., 2016. Interspecies differences in the metabolism of a multiester prodrug by carboxylesterases. J. Pharm. Sci. 105 (2), 989–995. https://doi.org/10.1002/jps.24632.

Gazin, V., Marsden, E., Marguerite, F., 2013. Oral propylparaben administration to juvenile male Wistar rats did not induce toxicity in reproductive organs. Toxicol. Sci. 136 (2), 392–401. https://doi.org/10.1093/toxsci/kft211. Studies performed in 2012 at Ricerca Biosciences.

Géniès, C., Jamin, E.L., Debrauwer, L., Zalko, D., Person, E.N., Eilstein, J., Grégoire, S., Schepky, A., Lange, D., Ellison, C., Roe, A., Salhi, S., Cubberley, R., Hewitt, N.J., Rothe, H., Klaric, M., Duplan, H., Jacques-Jamin, C., 2019. Comparison of the metabolism of 10 chemicals in human and pig skin explants. J. Appl. Toxicol. 39 (2), 385–397. https://doi.org/10.1002/jat.3730.

Golden, R., Gandy, J., Vollmer, G., 2005. A review of the endocrine activity of parabens and implications for potential risks to human health. Crit. Rev. Toxicol. 35 (5), 435–458.

Hewitt, N.J., Edwards, R.J., Fritsche, E., Goebel, C., Aeby, P., Scheel, J., Reisinger, K., et al., 2013. Use of human in vitro skin models for accurate and ethical risk assessment: metabolic considerations. Toxicol. Sci. 133 (2), 209–217. https://doi.org/10.1093/toxsci/kft080.

Hoberman, A.M., Schreur, D.K., Leazer, T., Daston, G.P., Carthew, P., Re, T., Loretz, L., Mann, P., 2008. Lack of effect of butylparaben and methylparaben on the reproductive system in male rats. Birth Defects Res. Part B Dev. Reproductive Toxicol. 83 (2), 123–133. https://doi.org/10.1002/bdrb.20153.

Hossaini, A., Larsen, J.J., Larsen, J.C., 2000. Lack of ooestrogenic effects of food preservatives (parabens) in uterotrophic assays. Food Chem. Toxicol. 38 (4), 319–323. https://doi.org/10.1016/s0278-6915(99)00160-x.

Hubbard, Troy, et al., Hubbard, T. D., Brix, A., Blystone, C. R, McIntyre, B. S., Shockley, K., Cunny, H., Waidyanatha, S., Turner, K. J., McBride, S., Roberts, G. K., 2020. Butylparaben multigenerational reproductive assessment by continuous breeding in Hsd:Sprague Dawley SD rats following dietary exposure. Reprod. Toxicol. 96, 258–272. https://doi.org/10.1016/j.reprotox.2020.07.006.

Janjua, N.R., Frederiksen, H., Skakkebaek, N.E., Wulf, H.C., Andersson, A.M., 2008. Urinary excretion of phthalates and paraben after repeated whole-body topical application in humans. Int. J. Androl. 31 (2), 118–130. https://doi.org/10.1111/j.1365-2605.2007.00841.x.

Kavlock, R.J., 1990. Structure-activity relationships in the developmental toxicity of substituted phenols: in vivo effects. Teratology 41 (1), 43–59. https://doi.org/10.1002/tera.1420410106.

Klimisch, H.J., Andreae, M., Tillmann, U., 1997. A systematic approach for evaluating the quality of experimental toxicological and ecotoxicological data. Regul. Toxicol. Pharmacol. 25 (1), 1–5. https://doi.org/10.1006/rtph.1996.1076.

Kolsek, K., Mavri, J., Sollner Dolenc, M., Gobec, S., Turk, S., 2014. Endocrine disruptome–an open source prediction tool for assessing endocrine disruption potential through nuclear receptor binding. J. Chem. Inf. Model. 54 (4), 1254–1267. https://doi.org/10.1021/ci400649p.

Kroes, R., Renwick, A.G., Feron, V., Galli, C.L., Gibney, M., Greim, H., Guy, R.H., Lhuguenot, J.C., van de Sandt, J.J., 2007. Application of the threshold of toxicological concern (TTC) to the safety evaluation of cosmetic ingredients. Food Chem. Toxicol. 45, 2533–2562.

Kuiper, G.G., Lemmen, J.G., Carlsson, B., et al., 1998. Interaction of estrogenic chemicals and phytoestrogens with estrogen receptor beta. Endocrinology 139 (10), 4252–4263. https://doi.org/10.1210/endo.139.10.6216. PMID: 9751507.

Laizure, S.C., Herring, V., Hu, Z., Witbrodt, K., Parker, R.B., 2013. The role of human carboxylesterases in drug metabolism: have we overlooked their importance? Pharmacotherapy 33 (2), 210–222. https://doi.org/10.1002/phar.1194.

Lemini, C., Jaimez, R., Avila, M.E., Franco, Y., Larrea, F., Lemus, A.E., 2003. In vivo and in vitro oestrogen bioactivities of alkyl parabens. Toxicol. Ind. Health 19 (2–6), 69–79. https://doi.org/10.1191/0748233703th177oa.

Lemini, C., Hernandez, A., Jaimez, R., Franco, Y., Avila, M.E., Castell, A., 2004. Morphometric analysis of mice uteri treated with the preservatives methyl, ethyl, propyl, and butylparaben. Toxicol. Ind. Health 20 (6–10), 123–132. https://doi.org/10.1191/0748233704th202oa.

Lester, C., Reis, A., Laufersweiler, M., Wu, S., Blackburn, K., 2018. Structure activity relationship (SAR) toxicological assessments: the role of expert judgment. Regul. Toxicol. Pharmacol. 92, 390–406. https://doi.org/10.1016/j.yrtph.2017.12.026.

Li, B., Sedlacek, M., Manoharan, I., Boopathy, R., Duysen, E.G., Masson, P., Lockridge, O., 2005. Butyrylcholinesterase, paraoxonase, and albumin esterase, but not carboxylesterase, are present in human plasma. Biochem. Pharmacol. 70 (11), 1673–1684. https://doi.org/10.1016/j.bcp.2005.09.002. Nov 25 2005.

Lipinski, C.A., Lombardo, F., Dominy, B.W., Feeney, P.J., 2001. Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. Adv. Drug Deliv. Rev. 46, 3–26.

Lipscomb, J.C., Fisher, J.W., Confer, P.D., Byczkowski, J.Z., 1998. *In vitro* to *in vivo* extrapolation for trichloroethylene metabolism in humans. Toxicol. Appl. Pharmacol. 152 (2), 376–387. https://doi.org/10.1006/taap.1998.8485.

Mahony, C., Bowtell, P., Huber, M., Kosemund, K., Pfuhler, S., Zhu, T., Barlow, S., McMillan, D.A., 2020. Threshold of toxicological concern (TTC) for botanicals - concentration data analysis of potentially genotoxic constituents to substantiate and extend the TTC approach to botanicals. Food Chem. Toxicol. 138, 111–182. https://doi.org/10.1016/j.fct.2020.111182. Epub 2020 Feb 10. PMID:32058013.

Manwaring, J., Rothe, H., Obringer, C., Foltz, D.J., Baker, T.R., Troutman, J.A., Hewitt, N.J., Goebel, C., 2015. Extrapolation of systemic bioavailability assessing skin absorption and epidermal and hepatic metabolism of aromatic amine hair dyes in vitro. Toxicol. Appl. Pharmacol. 287 (2), 139–148. https://doi.org/10.1016/j.taap.2015.05.016.

Ministry of Health and Welfare (MHW) Japan, 1997. Toxicity Testing Reports of Environmental Chemicals 5, vol. 257. Testing Laboratory: Hatano Research Institute, Japan. Food and Drug Safety Center, 729-5 Ochiai, Kanawaga.

Moos, R.K., Angerer, J., Dierkes, G., Bruning, T., Koch, H.M., 2016. Metabolism and elimination of methyl, iso- and N-butyl paraben in human urine after single oral dosage. Arch. Toxicol. 90 (11), 2699–2709. https://doi.org/10.1007/s00204-015-1636-0.

Munro, I.C., Ford, R., Kennepohl, E., Sprenger, J.G., 1996. Thresholds of toxicological concern based on structure-activity relationships. Drug Metab. Rev. 28 (1–2), 209–217. https://doi.org/10.3109/03602539608994000.

OECD, S.I.D.S., 1999. Initial Assessment Report: 4- Hydroxybenzoic Acid. Organization for Economic Cooperation and Development (OECD), UNEP Publications, 1999. http://www.inchem.org/documents/sids/sids/99967.pdf.

ECD 2004. Guideline *Test No. 428*: Skin Absorption: In Vitro Method. doi: https://doi.org/10.1787/9789264071087-en .

OECD, 2011. Guideline *Test No. 456: H295r Steroidogenesis Assay*. https://doi.org/10.1787/9789264122642-en, 2011.

OECD, 2012. Conceptual Framework for Testing and Assessment of Endocrine Disrupters *(as Revised in 2012)*.

OECD, 2016. Guideline *Test No. 455: Performance-Based Test Guideline for Stably Transfected Transactivation in Vitro Assays to Detect Oestrogen Receptor Agonists and Antagonists*. https://doi.org/10.1787/9789264265295-en, 2016.

OECD, 2018. The OECD QSAR Toolbox. http://www.oecd.org/chemicalsafety/oecd-qsar-toolbox.htm. (Accessed 19 January 2018).

OECD, 2020. ENV/JM/MONO(2020)16 - Series on Testing and Assessment No. 320: Case Study on the 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. http://www.oecd.org/officialdocuments/publicdisplaydocumentpdf/?cote=ENV/JM/MONO(2020)16&docLanguage=en.

OECD, 2021. OECD Series on Testing and Assessment, No. 331: Guidance Document on the Characterisation, Validation and Reporting of Physiologically Based Kinetic (PBK) Models for Regulatory Purposes. https://www.oecd.org/chemicalsafety/risk-assessment/guidance-document-on-the-characterisation-validation-and-reporting-of-physiologically-based-kinetic-models-for-regulatory-purposes.pdf.

Oishi, S., 2001. Effects of butylparaben on the male reproductive system in rats. Toxicol. Ind. Health 17 (1), 31–39. PubMed PMID: 12004923.

Oishi, S., 2002. Effects of propyl paraben on the male reproductive system. Food Chem. Toxicol. 40, 1807–1813.

Oishi, S., 2004. Lack of spermatotoxic effects of methyl and ethyl esters of p-hydroxybenzoic acid in rats. Food Chem. Toxicol. 42 (11), 1845–1849. PubMed PMID: 15350682.

Okubo, T., Yokoyama, Y., Kano, K., Kano, I., 2001. ER-dependent oestrogenic activity of parabens assessed by proliferation of human breast cancer MCF-7 cells and expression of ERalpha and PR. Food Chem. Toxicol. 39 (12), 1225–1232.

Ozaki, H., Sugihara, K., Watanabe, Y., Fujino, C., Uramaru, N., Sone, T., Ohta, S., Kitamura, S., 2013. Comparative study of the hydrolytic metabolism of methyl-, ethyl-, propyl-, butyl-, heptyl and dodecylparaben by microsomes of various rat and human tissues. Xenobiotica 43 (12), 1064–1072. https://doi.org/10.3109/00498254.2013.802059.

Routledge, E.J., Parker, J., Odum, J., Ashby, J., Sumpter, J.P., 1998. Some alkyl hydroxy benzoate preservatives (parabens) are oestrogenic. Toxicol. Appl. Pharmacol. 153 (1), 12–19. https://doi.org/10.1006/taap.1998.8544.

Sachana, M., 2019. An international effort to promote the regulatory use of PBK models based on non-animal data. Comput. Toxicol. 23–24. https://doi.org/10.1016/j.comtox.2019.01.002.

Scientific Committee on Consumer Products (SCCP), 2006. The Scientific Committee on Consumer Products (SCCP) Opinion on Parabens (Colipa N° P82), Adopted during the 9th Plenary Meeting of 10 October 2006. SCCP/1017/06. https://ec.europa.eu/health/ph_risk/committees/04_sccp/docs/sccp_o_074.pdf.

Scientific Committee on Consumer Safety (SCCS), 2013. Opinion on Parabens. Updated Request for a Scientific Opinion on Propyl- and Butylparaben Colipa N° P82. Adopted on 3 May 2013. SCCS/1514/13.

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

Shin, M.Y., Shin, C., Choi, J.W., Lee, J., Lee, S., Kim, S., 2019. Pharmacokinetic profile of propyl paraben in humans after oral administration. Environ. Int. 130, 104917. https://doi.org/10.1016/j.envint.2019.104917.

Soni, M.G., Carabin, I.G., Burdock, G.A., 2005. Safety assessment of esters of p-hydroxybenzoic acid (parabens). Food Chem. Toxicol. 43 (7), 985–1015. https://doi.org/10.1016/j.fct.2005.01.020.

Taketani, M., Shii, M., Ohura, K., Ninomiya, S., Imai, T., 2007. Carboxylesterase in the liver and small intestine of experimental animals and human. Life Sci. 81 (11), 924–932. https://doi.org/10.1016/j.lfs.2007.07.026.

Tozer, S.A., Kelly, S., O'Mahony, C., Daly, E.J., Nash, J.F., 2015. Aggregate exposure modelling of zinc pyrithione in rinse-off personal cleansing products using a person-orientated approach with market share refinement. Food Chem. Toxicol. 83, 103–110.

Tozer, S., O'Mahony, C., Hannah, J., O'Brien, J., Kelly, S., Kosemund-Meynen, K., Alexander-White, C., 2019. Aggregate exposure modelling of vitamin A from cosmetic products, diet and food supplements. Food Chem. Toxicol. 131, 110549.

Trott, O., Olson, A.J., 2010. Autodock vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. J. Comput. Chem. 31 (2), 455–461. https://doi.org/10.1002/jcc.21334.

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. FDA, 2018. Physiologically Based Pharmacokinetic Analyses — Format and Content Guidance for Industry.

Watanabe, Y., Kojima, H., Takeuchi, S., Uramaru, N., Ohta, S., Kitamura, S., 2013. Comparative study on transcriptional activity of 17 parabens mediated by oestrogen receptor α and β and androgen receptor. Food Chem. Toxicol. 57, 227–234. https://doi.org/10.1016/j.fct.2013.03.036.

WHO, I.P.C.S., 2005. International Programme on Chemical, Safety Workshop on Incorporating Uncertainty Variability into Risk, Assessment. Chemical-specific Adjustment Factors for Interspecies Differences and Human Variability:Guidance Document for Use of Data in Dose/Concentration-Response Assessment. World Health Organization, Geneva.

WHO, 2010. Characterization and Application of Physiologically Based Phamacokinetic Models in Risk Assessment. World Health Organization, Geneva.

Williams, F.M., 2008. Potential for metabolism locally in the skin of dermally absorbed compounds. Hum. Exp. Toxicol. 27, 277–280.

Wu, S., Blackburn, K., Amburgey, J., Jaworska, J., Federle, T., 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. 56 (1), 67–81. https://doi.org/10.1016/j.yrtph.2009.09.006.

Yang, C., Barlow, S.M., Muldoon Jacobs, K.L., Vitcheva, V., Boobis, A.R., Felter, S.P., Arvidson, K.B., Keller, D., Cronin, M.T.D., Enoch, S., Worth, A., Hollnagel, H.M., 2017. Thresholds of Toxicological Concern for cosmetics-related substances: new database, thresholds, and enrichment of chemical space. Food Chem. Toxicol. 109, 170–193.