Organisation for Economic Co-operation and Development

ENV/JM/MONO(2020)17

Unclassified English - Or. English

24 September 2020

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

Cancels & replaces the same document of 23 September 2020

Case Study on the use of Integrated Approaches for Testing and Assessment for Systemic Toxicity Arising from Cosmetic Exposure to Caffeine

Series on Testing and Assessment
No. 321

JT03465717

OECD Environment, Health and Safety Publications

Series on Testing and Assessment

No. 321

CASE STUDY ON THE USE OF INTEGRATED APPROACHES FOR TESTING AND ASSESSMENT FOR SYSTEMIC TOXICITY ARISING FROM COSMETIC EXPOSURE TO CAFFEINE

IOMC

INTER-ORGANIZATION PROGRAMME FOR THE SOUND MANAGEMENT OF CHEMICALS

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

Environment Directorate

ORGANISATION FOR ECONOMIC CO-OPERATION AND DEVELOPMENT

Paris 2020

About the OECD

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

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

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

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

This publication is available electronically, at no charge.

Also published in the Series on testing and Assessment link

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

or contact:

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

Forward

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

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

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

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

1. CASE STUDY ON USE OF AN INTEGRATED APPROACH TO TESTING AND ASSESSMENT (IATA) AND NEW APPROACH METHODS TO INFORM A THEORETICAL READ-ACROSS FORDERMAL EXPOSURE TO PROPYLPARABEN FROM COSMETICS, ENV/JM/MONO(2020)16.
2. CASE STUDY ON THE USE OF INTEGRATED APPROACHES FOR TESTING AND ASSESSMENT FOR 90-DAY RAT ORAL REPEATED-DOSE TOXICITY OF CHLOROBENZENE-RELATED CHEMICALS, ENV/JM/MONO(2020)18.
3. CASE STUDY ON THE USE OF INTEGRATED APPROACHES FOR TESTING AND ASSESSMENT TO INFORM READ-ACROSS OF P-ALKYLPHENOLS: REPEATED-DOSE TOXICITY, ENV/JM/MONO(2020)19.
4. CASE STUDY ON THE USE OF INTEGRATED APPROACHES TO TESTING AND ASSESSMENT FOR PREDICTION OF A 90 DAY REPEATED DOSE TOXICITY STUDY (OECD 408) FOR 2-ETHYLBUTYRIC ACID USING A READ-ACROSS APPROACH FROM OTHER BRANCHED CARBOXYLIC ACIDS, ENV/JM/MONO(2020)20.
5. CASE STUDY ON THE USE OF INTEGRATED APPROACHES TO TESTING AND ASSESSMENT FOR READ-ACROSS BASED FILLING OF DEVELOPMENTAL TOXICITY DATA GAP FOR METHYL HEXANOIC ACID, ENV/JM/MONO(2020)21.
6. CASE STUDY ON THE USE OF INTEGRATED APPROACHES TO TESTING AND ASSESSMENT FOR IDENTIFICATION AND CHARACTERISATION OF PARKINSONIAN HAZARD LIABILITY OF DEGUELIN BY AN AOP-BASED TESTING AND READ ACROSS APPROACH, ENV/JM/MONO(2020)22.

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

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

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

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

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

Abstract

This case study was developed to demonstrate how read-across can be applied in order to fill data gaps in an assessment of the potential risk for the consumer from exposure to caffeine. For this purpose, *in vivo* data from structural analogues have been used while assuming that no *in vivo* repeated dose toxicity data were available for the target substance caffeine.

For the read-across approach, several caffeine analogues have been identified by databases such as the QSAR toolbox v 4.3 and ToxCast. On the basis of an analysis of the pivotal mode of action (MOA) of caffeine and other methyl xanthines, theophylline appeared to be the most potent methyl xanthine.

Using a physiologically-based biokinetic (PBBK) model to estimate blood concentrations following exposures to caffeine in experimental animals and humans, a Margin of Internal Exposure (MoIE) of 25 was developed for caffeine based on *in vivo* data of theophylline. A MoIE of 25 is considered equivalent to the default Margin of Exposure (MOE) of 100 due to its greater precision for the chemical of concern because it is calculated as the ratio of a measure of internal exposure, such as blood concentration or target-tissue dose, rather than a measure of external exposure concentration or ingested dose.

The level of uncertainty was considered low since based on the considered common MoA between caffeine and the closest structural analogues, internal plasma levels derived from an acceptable *in vivo* animal study with theophylline were used as point of departure (NOAEL). The final Margin of Internal Exposure (MoIE) calculation was considered sufficiently conservative since it was based on estimated internal exposure with a PBBK model, rather than on external doses with inherent uncertainty related to route-to-route extrapolation and species/strain differences. Thus, calculation of internal exposures with a PBBK model can be used to replace the default uncertainty factor of 4 for interspecies differences in toxicokinetics. The case study demonstrates that we can successfully use read across supported by NAM to define a MoIE for caffeine based on theophylline data.

Table of Contents

Tables

Figures

1. Introduction

According to the European Regulation (EC) n°1223/2009, cosmetic products placed on the EU market must be safe. The manufacturer must ensure that they undergo an expert scientific safety assessment before they are put on the market. Caffeine is not used as preservative, colorant or UV filter, therefore, it is not listed in the annexes of the Cosmetic Products Regulation (CPR), but it has to undergo a safety assessment for the Cosmetic Safety Report (CSR).

Caffeine is a methyl xanthine compound (chemical name: 1,3,7-trimethylpurine-2,6-dione) and is metabolised in the liver mainly to three metabolites which have close structural, physicochemical and molecular similarity to caffeine. These are theophylline (1,3-dimethylxanthine), theobromine (3,7-dimethylxanthine), and paraxanthine (1,7-dimethylxanthine) (Berthou *et al*, 1989).

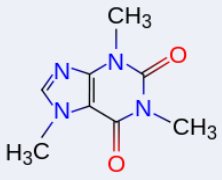
This case study was developed to demonstrate how read-across can be applied to fill data gaps in an assessment of the potential risk for the consumer from exposure to caffeine by using *in vivo* data from structural analogues while assuming that no *in vivo* repeated dose toxicity data were available for the target substance caffeine. In order to support the read-across approach and to decrease the level of uncertainty, a physiologically-based biokinetic (PBBK) model was used to simulate concentration – time profiles of caffeine in blood and liver, following a caffeine-specific ADME scenario which considered systemic exposure following potential aggregate exposure by both the dermal and oral route.

2. Purpose

2.1. Purpose of use

The purpose of this case study is to assess the potential risk from consumer exposure to caffeine by using a read-across approach which utilised *in vivo* data from close structural analogues, paraxanthine, theobromine and theophylline, while assuming that no *in vivo* repeated dose toxicity data were available for caffeine. The exposure assessment covered aggregate exposure from dermal (cosmetics) and oral (food/drink) exposure. In order to increase the confidence into the read-across approach, *in silico*/*in vitro* biokinetic and toxicodynamic data were generated and a physiologically-based biokinetic (PBBK) model was established to enable a robust estimate of the internal concentration of caffeine after both dermal and oral exposure. Based on the toxicity data for the metabolites of caffeine which were used as structural analogues for read-across, a no-observed-adverse-effect-level (NOAEL) modelled plasma concentration was derived which was compared with the modelled plasma concentrations resulting from human dermal and oral exposure to achieve a Margin of Internal Exposure (MoIE) value for the safety assessment.

2.2. Target chemical(s)/category definition

Name	Caffeine
IUPAC Name	1,3,7-trimethylpurine-2,6-dione
CAS #	58-08-2
Chemical formula	$C_8H_{10}N_4O_2$
SMILES	CN1C=NC2=C1C(=O)N(C(=O)N2C)C
2D structure	

For the purpose of this case study, we assessed the pure chemical and did not consider any impurities.

2.3. Endpoint(s)

The endpoint of interest are data on systemic toxicity for caffeine which may enable to derive an adequate NOAEL to be used for the risk assessment.

2.4. Exposure information

Caffeine is a naturally occurring compound found in plants. Exposure to caffeine may occur from different sources, i.e. through food, medicines and consumer products. In cosmetics and personal care products, caffeine is used mainly in hair care, skin care, makeup and hygiene products. The commercially available formulations of caffeine normally contain 2% caffeine. The pivotal exposure to caffeine from cosmetic products is by the dermal route.

An aggregate exposure assessment may take into account all relevant sources including food, drink and cosmetic products and all relevant routes of exposure such as oral and dermal.

In Table 1, the daily exposure levels are listed for the different cosmetic product categories in Europe which may contain caffeine. According to the available data, caffeine may be used in most cosmetic product categories, but not in oral care.

Using a deterministic exposure assessment as outlined in the SCCS Notes of Guidance (2018), a worst case exposure estimate from use of cosmetic products including shower gel, shampoo, hair styling products, body lotion, face cream, hand cream, liquid foundation, lipstick, and deodorant/antiperspirant, resulted in a potential maximum external exposure of 4.6 mg/kg/d. This value is based on the traditional approach to modelling exposure assuming that all cosmetic products contain caffeine at a use concentration of 2% (230.96 mg/kg bw x 0.02) and are all used by all persons at a high amount per use, and at a high frequency per day (SCCS, 2018). Moreover, the aggregate exposure is based on the summation of all individual product exposures thus leading to a very conservative calculation.

Table 1. Daily exposure levels for different cosmetic product categories in Europe which may contain caffeine, calculated by multiplying daily amounts and retention factor (SCCS, 2018)

Product Type	*Estimated daily amount applied (g/d)*	*Relative daily amount applied (mg/kg bw/d)[1]*	*Retention factor[2]*	*Calculated daily exposure (g/d)*	*Calculated relative daily exposure (mg/kg bw/d)[1]*
Bathing, showering					
Shower gel	18.67	279.20	0.01	0.19	2.79
Hair care					
Shampoo	10.46	150.49	0.01	0.11	1.51
Hair styling products	4.00	57.40	0.10	0.40	5.74
Skin care					
Body lotion	7.82	123.20	1.00	7.82	123.20
Face cream	1.54	24.14	1.00	1.54	24.14
Hand cream	2.16	32.70	1.00	2.16	32.70
Make-up					
Liquid foundation	0.51	7.90	1.00	0.51	7.90
Lipstick, lip salve	0.057	0.90	1.00	0.057	0.90
Deodorant					
Deodorant non- spray	1.50	22.08	1.00	1.50	22.08
Deodorant spray	0.69	10.00	1.00	0.69	10.00
Aggregate exposure					230.96

[1]The specific body weight of the persons involved in the study is used and not the default value of 60 kg
[2]The retention factor was introduced to take into account rinsing off and dilution of finished products by application on wet skin or hair (e.g. shower gels, shampoos)

Thus, the worst-case exposure estimate of 4.6 mg/kg/d (230.96 mg/kg bw x 0.02) from use of cosmetic products according to a deterministic exposure assessment was taken into account for the safety assessment.

Caffeine exposure from other sources includes oral intake from coffee, tea, energy drinks, cola and chocolate (EFSA, 2015). The worst-case (maximum 95th percentile) of caffeine intake from all food/drink sources for all days is estimated to be 648 mg/person/d (10.8 mg/kg bw) for adults (18 to <65 yr) and 786 mg/person/d (13.1 mg/kg bw) for elderly (65 to <75 yr, ≥75 yr). The latter was used for the further assessment.

Caffeine exposure may occur from different sources, i.e. through food, medicines and consumer products. In this case study, we used for the assessment the sum of the exposure via the oral route:

1. *Worst case (maximum 95th percentile) of caffeine intake from all food/drink sources for all days for the most exposed population (13.1 mg/kg bw)*
2. *Worst case exposure estimate of 4.6 mg/kg/d from use of cosmetic products according to a deterministic exposure assessment*

3. Hypothesis for the analogue approach

Our hypothesis, which is supported by extensive data published in the scientific literature (CIR, 2018; Hawke *et al*, 2000; Monteiro *et al*, 2016; Muller & Jacobson, 2011), is that we can use repeated dose toxicity data from an analogue of caffeine to support the risk assessment. Caffeine is metabolised rapidly and extensively in the body to predominantly three metabolites, i.e. paraxanthine (84%), theobromine (12%) and theophylline (4%) which were found to be chemically similar analogues of caffeine based on structure, phys-chem, ADME data and *in vitro* safety data determined using ToxCast assays. The availability of reliable repeated dose data was also a factor to choose the source substance. The NOAEL from a rabbit prenatal developmental toxicity study on theophylline which proved to be the most appropriate analogue based on potency and supportive data (e.g. ToxCast, see section 5.2.3for details), was then used for the safety assessment of caffeine, also because internal plasma levels were available for this animal study.

3.1. Chemical identity and composition

Relevant physicochemical data are summarised in Table 2.

Table 2. Comparison of physico-chemical and molecular properties (measured and predicted values)

	Target Substance	Analogue 1	Analogue 2	Analogue 3
Name	Caffeine	Theophylline (4%)	Theobromine (12%)	Paraxanthine (84%)
Molecular Weight	194.194 g/mol **194.2 g/mol**	180.167 g/mol **180.2 g/mol**	180.167 g/mol **180.2 g/mol**	180.167 g/mol **180.2 g/mol**
Log P_{ow}	-0.07 **0.0**	-0.02 **0.0**	-0.78 **-0.7**	-0.63 **-0.3**
Vapour Pressure	$9.0X10^{-7}$ mm Hg at 25 °C (est), 0.00011999 Pa **non volatile**	$5.12X10^{-9}$ mm Hg at 25 °C (est), 0.00006826 Pa **non volatile**	$1.13X10^{-11}$ mm Hg at 25 °C (est), 0.00000015 Pa **non volatile**	 **non volatile**
Water Solubility	22000 mg/ L (at 25 °C)	8300 mg/ L (at 25 °C)	3300 mg/ L (at 25 °C)	1000 mg/L (at 25°C)
Lipinski rule	**bioavailable**	**bioavailable**	**bioavailable**	**bioavailable**
Living skin + RF absorption	**Very high**	**Very high**	**Very high**	**Very high**
Tanimoto coefficients, based on MACCS fingerprints	**1.0 (target)**	**0.95**	**0.90**	**0.90**

Source: PubChem

The comparison of physico-chemical and molecular properties shows that:

1. *Caffeine and its three analogues theophylline, theobromine and paraxanthine have comparable physico-chemical properties (similar MW and logP range, similar negligible volatility, sparingly to slightly water soluble)*
2. *According to the Lipinski rule, all chemicals are predicted to be bioavailable and have a good skin penetration*
3. The structural similarity between the target compound and the analogues is very high (Tanimoto score >0.9)

3.2. Route and duration of exposure

The external exposure values in the risk assessment have been presented in section 2.4 Exposure information. The use duration for the different cosmetic products is estimated to be daily for in general one to two times per day.

3.3. Kinetics

3.3.1. Skin penetration

The dermal penetration of caffeine has been studied extensively *in vitro* and *in vivo*. On the basis of the available studies, the skin penetration was in the range of 1.1% – 40% (1.25±0.17% - 41±20.03%) *in vitro* and 2.5% – 62% (2.5 - 61.8±5.4%) *in vivo*. The skin penetration rates are depending on various factors such as the use of vehicle/formulation, test concentration, site of the skin used, blocking of hair follicles, test duration and occlusion. Since the *in vivo* data in humans resulting in 67% skin penetration were performed with acetone as vehicle which has an impact on the skin penetration, an average value of 50% was selected for the safety assessment on the basis of the *in vitro* skin penetration data performed with phosphate-buffered saline (PBS) (Hewitt *et al*, 2019) (Table 3, Figure 4).

A permeability coefficient (Kp) is also needed in the PBBK model to calculate a rate of absorption. Doucet *et al.* (1998) reported Kp measured using two different oil/water vehicles, with an average of $4.1x10^{-4}$ cm/h. This value is close to the value $2.1x10^{-4}$ cm/h measured by Dias *et al.* (1999) using a saturated solution of caffeine in water.

Table 3. Summary of Skin Penetration Studies with Caffeine (*in vitro* and *in vivo*)

Study Type	Caffeine concentration	Vehicle / Formulation	Fraction absorbed	Reference
In vitro (human skin)	1.08 µg/cm²	0.01M phosphate-buffered saline (PBS)	41±20.03% (mass balance: 97.11±2.21%)	Hewitt *et al*, 2019 (in press)
In vitro (human skin, 24 h)	1% (w/w), 260 mg/cm²	W/O/W, O/W	1.25±0.17% (W/O/W) 3.21±0.18% (O/W)	Doucet *et al*, 1998
In vitro (human skin, 24 h)	250 µg/cm²	Ethanol (70%)	17% (unblocked hair follicles), 7% (blocked hair follicles)	Trauer *et al*, 2009
In vitro (human skin, 72 h)	4 mg/mL, 10 µL/cm², repeated dosing (24 and 48 h)	ethanol:water (1:1, v/v)	Dermal delivery 13.69±5.95%	Toner *et al*, 2009
In vitro (human skin, 24 h)	1% (w/w), 10 µg/cm²	O/W, W/O	15-20%	Luo & Lane, 2015
n vivo Human subjects			2.5 – 3.3%	Zesch *et al* 1979
In vivo Human subjects (24 h), blocked and unblocked hair follicles	2.5%, 10 µg/cm²	Ethanol and propylene glycol	Hair follicles unblocked: 24.9±1.05% (receptor fluid) Hair follicles blocked: 12.4±0.9% (receptor fluid)	Otberg *et al*, 2008 Trauer *et al*, 2009
In vivo Human subjects (ventral forearm)	4 µg/cm²	acetone	About 50%	Luo & Lane, 2015
In vivo Human subjects (forearm, non-occlusive patch, 24 h)	4 µg/2.5cm²	acetone	32.1±4.2% (22-40 yr) 61.8±5.4% (65-86 yr)	Luo & Lane, 2015

No human *in vitro* or *in vivo* dermal delivery data for theobromine or paraxanthine could be found. For theophylline a study in human *in vitro* skin reported a Kp value of 2,1x10E-

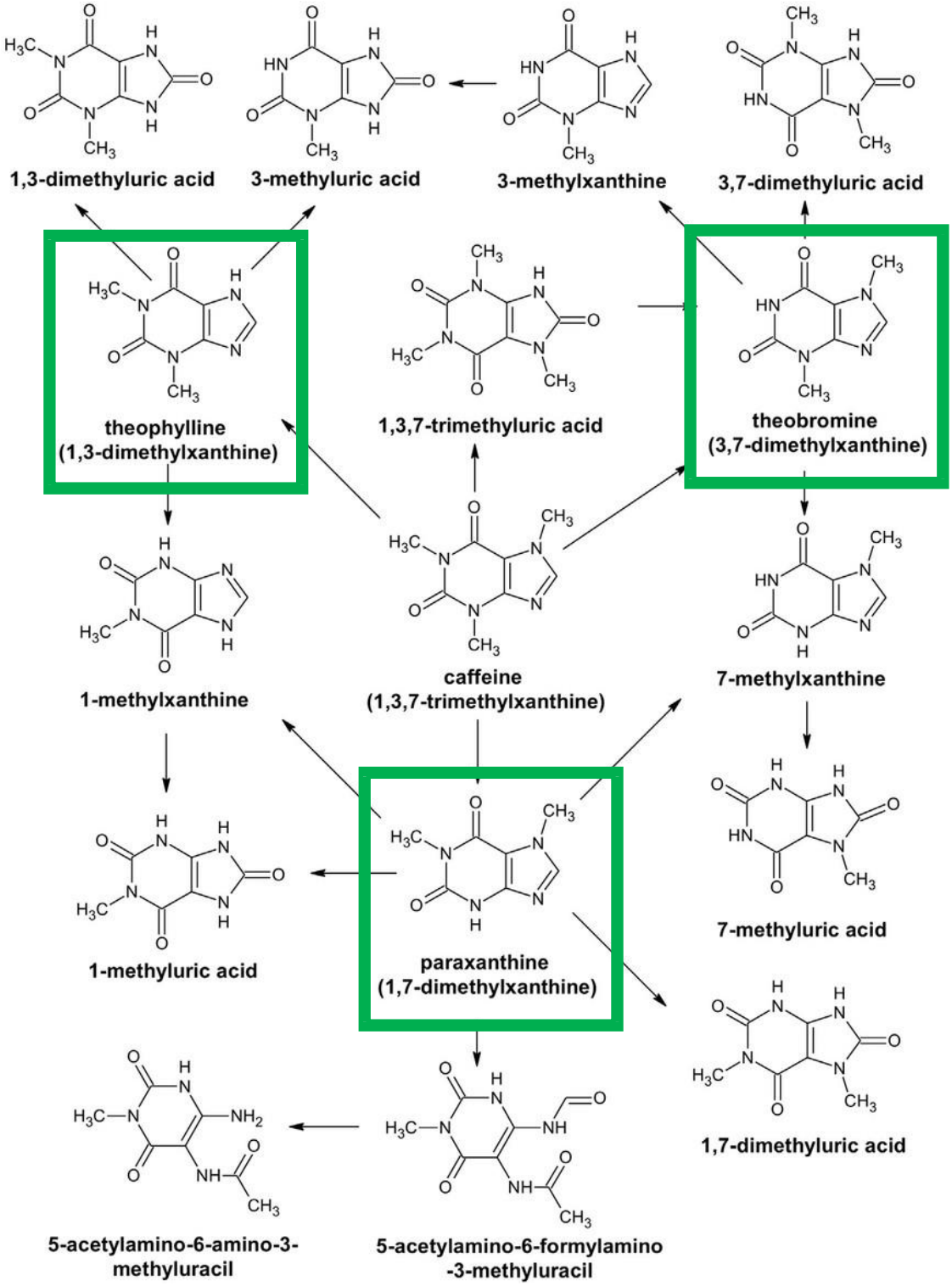
05 (Kopecna *et al*, 2017). Skin metabolism of caffeine could not be observed (Gajewska *et al*, 2014 & 2015, Géniès *et al*, 2019).

3.3.2. Metabolism

Once caffeine has penetrated the skin and becomes systemically available, it is metabolised in the liver by CYP450 1A2 mainly to three metabolites, i.e. theophylline (1,3-dimethylxanthine), theobromine (3,7-dimethylxanthine), and paraxanthine (1,7-dimethylxanthine) (see Figure 1), corresponding respectively to 4%, 10-12% and 80-84% of the parent caffeine (Puchem and DrugBank). Metabolism in the skin is regarded as negligible for caffeine since skin expresses little amount of CYP450 1A2 (Gajewska *et al*, 2014 & 2015, Géniès *et al* 2019). Caffeine, theophylline and theobromine are metabolically converted into one another in the liver where ca. 90-99 % of the metabolism takes place.

Following oral intake, caffeine, theophylline, theobromine and paraxanthine are readily and completely absorbed. The biotransformation is catalysed by the cytochrome P-450 monooxygenase system and leads to the formation of demethylated compounds or, through oxidation, to the urate and/or hydration to the diaminouracil.

Figure 1. The metabolites of caffeine (adapted from Gracia-Lor *et al.* (2017), major metabolites are highlighted)



3.3.3. Prediction of in vitro/in vivo caffeine metabolism in humans

In vitro models measuring the metabolism of caffeine are difficult to optimise but all investigations indicate that this is a low clearance compound (Table 4). Based on results by Berthou *et al.* (1989), the biotransformation of caffeine is comparable in human liver slices, microsomes and hepatocyte cultures. The reason for the low metabolic rate *in vitro* is not due to the lack of uptake into cells since (a) metabolism is also low in subcellular fractions, and (b) permeability across Caco-2 cells is high (P_{app} = 28 x10^{-6} cm/s cf testosterone P_{app} = 17 x10^{-6} cm/s [Cosmetics Europe data (2017)]). Cosmetics Europe (2017) experiments investigating the *in vitro* biokinetics of caffeine and its metabolites indicate that very little of the parent compound is detected in the cells; whereas, the metabolites appear in both the cells and medium.

Table 4 shows the different conditions under which caffeine metabolism has been tested. In general, when unlabelled compound was used in low concentrations (up to 200 μM) or the duration of incubation was short (3 to 8 h), there was no depletion of parent chemical and no production of any metabolites (analysed by UV-HPLC or LC-MS). In order to detect metabolism, experiments were adapted to incorporate radiolabelled (Berthou *et al*, 1989) caffeine at concentrations markedly higher (at least 1 mM) than those present *in vivo* (50-100 μM) and with extended incubation durations (up to 72 h).

Incubations with HepaRG cells required them to be in 3D spheroid form, rather than in 2D monolayers, to observe sufficient metabolism. Shibata *et al.* (2000) only used 10 μM caffeine and a 2 h incubation; however, the C_{Lint} value derived was lower than that of other groups who used durations up to 72 h, suggesting that this incubation may not have detected sufficient parent compound depletion to make an accurate calculation of the C_{Lint}. To derive K_m and V_{max} kinetics values, Berthou *et al.* (1989) tested concentrations up to 50 mM (which is well above the K_m for the CYPs involved). The Eadie-Hofstee plot from incubations with human liver microsomes showed a biphasic nature of caffeine transformation, indicating that two distinct enzymes with different substrate affinities were involved in this reaction.

Berthou *et al.* (1989) calculated a half-life ($t_{1/2}$) of hepatic elimination for a normal average drug intake of 1400 μM (~270 mg of caffeine) to be 4.5 h, assuming that the caffeine metabolism is linear with time in hepatocyte cultures. This value is in agreement with the $t_{1/2}$ determined *in vivo* (Bonati and Garattini 1984).

Table 4. Assay conditions and results from *in vitro* assays measuring caffeine metabolism.

Model Type	Caffeine concentration	Volume	Protein/cell/ spheroid concentration	Incubation time	Result	Reference
Human liver S9	5 μM	1.2 mL	2 mg/mL	4 h	No parent depletion and no metabolites formed (LC-MS)	Eilstein *et al*, 2020
Primary human hepatocyte suspensions	5 μM	0.25 mL	0.2 million cells (0.8 million/mL)	3 h	No parent depletion and no metabolites formed (LC- MS)	Eilstein *et al*, 2020
HepaRG monolayers (no overlay)	50 μM	1 mL	Confluent in 24-well plates	8 h	No parent depletion and no metabolites formed (UV- HPLC)	Cosmetics Europe (2017)
HepaRG sandwich monolayers (Matrigel/coll agen)	50 μM	1 mL	Confluent in 24-well plates	8 h	No parent depletion and no metabolites formed (UV- HPLC)	Cosmetics Europe (2017)
HepaRG spheroids	50 μM	1 mL	24 spheroids/we ll (24-well plates)	8 h	No parent depletion and no metabolites formed (UV- HPLC)	Cosmetics Europe (2017)
HepaRG monolayers (no overlay)	200 & 1000 μM	0.5 mL	Confluent in 24-well plates	72 h	No parent depletion, no metabolites formed (UV- HPLC, extraction protocol – 10x concentrated)	Cosmetics Europe (2017)
HepaRG spheroids	1000 μM	0.6 mL	48 spheroids/we ll (24-well plates)	72 h	No parent depletion but 3 metabolites formed (UV-HPLC, extraction protocol – 10x concentrated). Theobromine conc in medium 5x paraxanthine conc (as expected from lit). Theo-bromine and paraxanthine metabolites at equivalent conc in spheroid extracts.	Cosmetics Europe (2017)
Human liver microsomes	12 & 1000 mM ^{14}C-caffine for metabolite ID; 0.05-50 mM for kinetics	3 mL	3 mg (1 mg/mL)	0.5 h	Theobromine, paraxanthine and theophylline produced, formation linear over 60 min. Low affinity site: K_{m2} = 21.8 ± 2 mM, V_{max} = 860 pmole / min / mg protein High affinity site: K_{m1} = 1.32 ± 0.16 mM, V_{max} = 101 ± 14 pmole/min/mg protein. Mean rate of metabolism = 174 ± 116 pmole/min/mg protein.	Berthou *et al.*, 1989

The prediction of *in vivo* caffeine metabolism has also been conducted by incorporating data from recombinant enzymes (Ginsberg *et al*, 2004). The metabolism of caffeine by CYP is almost completely attributed to CYP1A2 (Table 5). The K_m and V_{max} kinetics from these experiments, together with the relative activity factor (which takes into account the relative abundance of each CYP in the liver), can be used to predict the formation of each metabolite.

Table 5. Caffeine Metabolism to Paraxanthine, Theobromine and Theophylline (Ginsberg *et al*, 2004)

	Paraxanthine (1,7X)			Theobromine (3,7X)			Theophylline (1,3X)		
CYP Isoform	V_{max}[1]	K_m[2]	CL_{int} V_{max}/K_m	V_{max}[1]	K_m[2]	CL_{int} V_{max}/K_m	V_{max}[1]	K_m[2]	CL_{int} V_{max}/K_m
1A1	2.69	0.59	4.56	0.82	0.41	2.0	Nd[3]	Nd[3]	Nd[3]
1A2	30.5	0.19	161	3.0	0.16	18.8	1.12	0.25	4.5
2E1	Nd[3]	Nd[3]	Nd[3]	0.48	1.44	0.33	0.36	0.84	0.43
3A4	Nd[3]	Nd[3]	Nd[3]	Nd[3]	Nd[3]	Nd[3]	Nd[3]	Nd[3]	Nd[3]

[1]V_{max} units in moles of metabolite formed per h per mole CYP.
[2]K_m units in mmol/liter.
[3]not detectable

The comparison of kinetic properties shows that:

1. *Skin penetration data for caffeine show variability based on solvent and assay conditions. An average value of 50% was selected for the safety assessment from the in vitro skin penetration data performed with PBS. Skin penetration data for the analogues are scarce, however, due to structural similarity and the study of Kopecna et al. (2017) we assume similar penetration rate for theophylline.*
2. *Caffeine is not metabolised by the skin. In vitro studies using hepatic models indicate that caffeine is a low clearance compound i.e. slowly metabolised.*

3.4. Structural profiling and available data for hazard identification of caffeine, including possible endocrine activity

The OECD QSAR toolbox v4.3 was searched for *in silico* alerts on different toxicological endpoints (Annex 2).

3.4.1. Skin and eye irritation

No *in silico* alerts (OECD QSAR toolbox v4.3, Annex 2)

3.4.2. Skin sensitisation

No *in silico* alerts (OECD QSAR toolbox v4.3, Annex 2)

- Protein binding – none
- DPRA assay: Cys / Lys – no reaction

3.4.3. Genotoxicity

No *in silico* alerts found for *in vitro* mutagenicity (Ames) and DNA alerts for CA and MNT according to OECD QSAR toolbox v4.3 (Annex 2) and Chemtunes v1.2. An *in silico* alert for micronucleus formation *in vivo* (H-acceptor-path3-H-acceptor) was found in OECD QSAR toolbox v4.3.

Multiple *in vitro* genotoxicity tests were performed with caffeine and yielded conflicting results. The majority of bacterial *in vitro* tests yielded positive results, the majority of mammalian cell *in vitro* genotoxicity assays yielded negative results (CIR, 2018). However, positive results were only observed in *in vitro* studies without metabolic activation (those *in vitro* studies with metabolic activation were negative). Relevant *in vivo* genotoxicity studies involving the metabolite theophylline yielded negative results. Carcinogenicity studies performed by the NTP (1998) with the metabolite theophylline revealed negative results, further eliminating the need for concern of the positive genotoxicity studies (CIR, 2018).

3.4.4. Potential endocrine activity

In silico screening tools were used in order to look for alerts on potential endocrine activity with the focus on estrogenic, androgenic, thyroidal and steroidogenic activities (EATS) (EFSA/ECHA Guidance, 2018).

- OECD QSAR toolbox v4.3, Annex 2 (oestrogenic activity)

- Caffeine does not bind to the oestrogen receptor

- Endocrine Disruptome (http://endocrinedisruptome.ki.si/) (oestrogenic, androgenic, thyroidal, steroidogenic activity; the endocrine action of a molecule can be predicted against 16 structures, belonging to 12 nuclear receptors)
  - Between 25 and 50% probability to act as an androgen receptor antagonist
  - Less than 25% probability to bind to the androgen, oestrogen α and β, glucocorticoid, liver X α and β, PPAR α β ϒ, RXR α, thyroid α and β receptor
- VEGA platform (estrogenic activity; VEGA provides tens of QSAR models to predict tox, ecotox, environ, and phys-chem properties of chemical substances) (https://www.vegahub.eu/download/)
  - Receptor relative binding affinity inactive
  - Receptor-mediated effect non-active

The screening included the analogues theobromine, theophylline and paraxanthine which also showed only low *in silico* EATS activities (see Table 6).

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. Therefore, Endocrine Disruptome has a very large applicability domain while providing semi-quantitative predictions. These properties, together with the possibility of inspecting docked poses, makes it a more insightful tool than other QSAR models that usually simply discriminate between binders and nonbinders. These docking simulations were used to characterise the binding propensities of caffeine and its analogues 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. 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. These thresholds were established according to a conservative approach since Kolšek and co-authors (2014) decided that the true-positive rate was a more important than the true-negative rate for the division of the probability classes. The arithmetic mean of the five docking scores was retained as the final score for the quantitative description of the binding affinities of chemicals. These final scores were then compared to critical score thresholds (specific for each receptor) and associated with color-coded binding probability classes: green, yellow, orange and red. These colours indicate low, low intermediate, high intermediate and high binding probabilities, respectively.

The docking simulation results are in Table 6. They show that all four analogues are associated with a low binding probability class (green colour). The only exception to this tendency is represented by the androgen receptor (AR) in antagonistic conformation (AR an.) which is associated with the lower class of intermediate binding probability (yellow colour).

Overall, the available *in silico* data indicate either a lack of or very low EATS receptor-mediated activities. Because of the absence of *in silico* alerts, no hypothesis could be derived from these results to pursue additional generation of experimental *in vitro* data on potential endocrine activity. Moreover, there are several tests in the ToxCast database showing weak *in vitro* endocrine activity for caffeine. However, these tests are mere screening assays, no GLP or OECD guideline studies but at the most OECD Conceptual Framework Level 2 studies. Many tests are only borderline active, often only the highest concentration is above baseline, or it is stated that the result is potentially confounded by overfitting.

From the large epidemiological database EFSA has concluded that no health concerns are to be expected from single doses and habitual caffeine intakes up to 200 mg (about 3 mg/kg bw for a 70-kg adult) consumed by pregnant women do not give rise to safety concerns. Also 'no health concerns in relation to male fertility have been raised by other bodies in previous assessments for this level of habitual caffeine consumption and no new data have become available on these or other clinical outcomes which could justify modifying these conclusions' (EFSA, 2015).

Table 6. Summary of *in silico* screening results for EATS activities

Endocrine Disruptome				
	Caffeine	**Paraxanthine**	**Theobromine**	**Theophylline**
Androgen receptor				
Androgen receptor (antag)				
Oestrogen receptor α				
Oestrogen receptor α (antag)				
Oestrogen receptor β				
Oestrogen receptor β (antag)				
Glucocorticoid receptor				
Glucocorticoid receptor (antag)				
Liver X receptor α				
Liver X receptor β				
PPAR α				
PPAR β				
PPAR γ				
RXR a				
Thyroid receptor α				
Thyroid receptor β				
OECD QSAR Toolbox				
Oestrogen receptor binding	Non-binder	Non-binder	Non-binder	Non-binder
VEGA				
Receptor relative binding affinity	Inactive	Inactive	Inactive	Inactive
Receptor mediated-effect	Non-active	Non-active	Non-active	Non-active

■: Probability of acting as NON-BINDER = > 0.75
■: Probability of acting as NON-BINDER = 0.50 - 0.75

The available in silico data from OECD QSAR toolbox, VEGA and the Endocrine Disruptome indicate either a lack of or very low EATS receptor-mediated activities. The predictions are the same for caffeine and the analogues supporting the suitability of the

analogue selection. These results do not support an endocrine mode of action for the source or target chemicals.

3.5. Mode of Action

The hypothesis is that caffeine and its analogues paraxanthine, theobromine and theophylline have a similar mode of action across species, but with different potencies as indicated below. Therefore, any adequate *in vivo* data from these analogues may be used in order to fill the data gaps of caffeine.

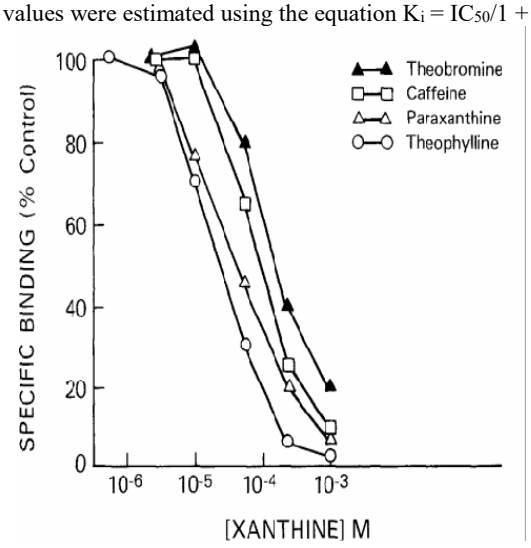
Several different mechanisms are proposed to mediate the pharmacological methyl xanthine activity at the cellular level, i.e. antagonism of adenosine receptors, phosphodiesterase inhibition, modulation of GABA receptor action, and regulation of intracellular calcium levels (EFSA, 2015; Monteiro *et al*, 2016). The pivotal mechanism of action of methyl xanthines is considered to involve a non-selective blocking of A1- and A2-adenosine receptors, thereby competitively inhibiting the action of adenosine in the cells.

Theophylline appears to be the most potent methyl xanthine according to the inhibitory constant (K_i) for the A1-adenosine receptor in rat brain and is therefore considered to be the suitable compound for read- across (Daly *et al*, 1983, Figure 2). The specific relative potency factors (RPF) are calculated as follows:

- K_i Theophylline – Adenosine receptor 12 µM – RPF 1
- K_i Paraxanthine – Adenosine receptor 30 µM – RPF 0.40 (1:2.5)
- K_i Caffeine – Adenosine receptor 50 µM – RPF 0.24 (1:4.16)
- K_i Theobromine – Adenosine receptor 120 µM – RPF 0.10 (1:10)

Figure 2. Effect of theophylline, paraxanthine, caffeine and theobromine on A1-Adenosine Receptor Systems in Rat Brain

Inhibition of binding of [^{3}H]-cyclohexyladenosine to rat cerebral cortical membranes by methyl xanthines. Membranes were incubated with 1 nM[^{3}H]N^6- cyclohexyladenosine in the absence or presence of various methyl xanthines. IC_{50} values are means of three determinations. The K_i values were estimated using the equation $K_i = IC_{50}/1 + [agonist]/EC_{50}$ (Daly *et al*, 1983)



Caffeine and other methyl xanthines affect the cardiovascular system, central nervous system, body temperature and hydration status via the mechanisms described above. In pregnant women, they can cause increase in maternal levels of 3'5'-cyclic adenosine monophosphate and epinephrine, which could lead to uteroplacental vasoconstriction and decreased intervillous placental blood flow, which could restrict foetal growth. Another hypothesis is that caffeine inhibits phosphodiesterase, leading to an increase in cellular cyclic adenosine monophosphate, which may interfere with foetal growth (Sengpiel *et al.* 2013).

The mechanism of action of methyl xanthine at the cellular level is caused by antagonism of adenosine receptors, phosphodiesterase inhibition, modulation of GABA receptor action, and regulation of intracellular calcium levels.

Theophylline appears to be the most potent methyl xanthine according to the inhibitory constant (K_i) for the A1-adenosine receptor in rat brain and is therefore considered to be the suitable compound for read-across, representing the worst-case scenario.

Possible MOA of caffeine and other methyl xanthines on the developing foetus is the uteroplacental vasoconstriction and decreased intervillous placental blood flow, which could restrict foetal growth

4. Source chemicals

4.1. Identification and selection of source chemicals

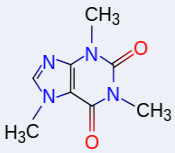
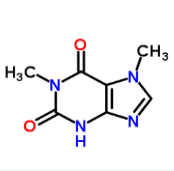
An initial screening of the structural analogues was performed with the CE-ToxGPS software that resulted in the identification of 70 analogues that had a structural similarity of above 70% based on the Tanimoto score (Annex 1). The next step was the identification of structural analogues with relevant *in vivo* repeated dose toxicity data. The literature search revealed repeat-dose and/or reproductive toxicity data in experimental animals for paraxanthine (1,7-dimethylxanthine), theobromine (3,7-dimethylxanthine) and theophylline (1,3-dimethylxanthine) (York *et al*, 1986; Theocorp Holding Company, 2010; OECD, 2001) (Table 8). The OECD QSAR Toolbox v4.3 has been used to profile caffeine and the three analogues (Annex 2).

Given the fact that caffeine, theophylline, theobromine and paraxanthine have the same mode of action (antagonism of adenosine receptors) they are considered as suitable analogues. The relative potency of the four chemicals can be derived from the inhibitory constant. Additionally, at least 94% of the parent caffeine is metabolised across these three metabolites (4%, 10% and 80% respectively for theophylline, theobromine and paraxanthine (PubChem, Drub Bank) which justifies the focus on these three analogues as source chemicals.

4.2. List of source chemicals

The common chemical identifiers (including CAS number, name and composition) and chemical structures of caffeine and its three major metabolites are summarised in Table 7.

Table 7. Comparison of substance identification, structure and chemical classifications

	Target Substance	Analogue 1	Analogue 2	Analogue 3
Name	Caffeine	Theophylline	Theobromine	Paraxanthine
CAS No:	58-08-2	58-55-9	83-67-0	611-59-6
2D Structure				
Molecular Formula	$C_8H_{10}N_4O_2$	$C_7H_8N_4O_2$	$C_7H_8N_4O_2$	$C_7H_8N_4O_2$
Chemical Class	Purine	Purine	Purine	Purine
Chemical Sub-Class	Methyl xanthine	Methyl xanthine	Methyl xanthine	Methyl xanthine
Chemical Sub-Class	1,3,7-trimethyl xanthine	1,3-dimethyl xanthine	3,7-dimethyl xanthine	1,7-dimethyl xanthine

5. Justification of data gap filling

5.1. Data gathering

The following sources were used to retrieve the necessary data: scientific literature (see references), COSMOS database, PubChem, *in silico* profiles, OECD Toolbox, and Cosmetics Europe data (2017).

5.2. Data and methods

5.2.1. Identification of structural analogues with suitable repeat-dose data

Already presented under section 4.1.

5.2.2. Repeat-dose toxicity

The NOAEL and LOAEL (lowest-observed-adverse-effect-level) derived from relevant repeat-dose and reproductive toxicity data in experimental animals are summarised in Table 8 for theophylline, theobromine and paraxanthine. Detailed study summaries are included in Annex 3.

Table 8. Repeat-dose and reproductive toxicity data for theophylline, theobromine and paraxanthine (Klimisch score 1 or 2)

In vivo data	Theophylline	Theobromine	Paraxanthine
Repeat-dose toxicity	Feeding study, rat, 14 weeks, 0, 1000, 2000 and 4000 ppm (m: ca. 75, 125, 250 mg/kg bw/d; f: ca. 75, 125, and 275 mg/kg bw/d), LOAEL: 75 mg/kg bw/d (m: nephropathy, f: periarteritis) (NTP, 1998) Oral (gavage) study, rat, 14 weeks, 0, 37.5, 75, and 150 mg/kg bw/d, LOAEL: 37.5 mg/kg bw/d (periarteritis) (NTP, 1998) Feeding study, mouse, 14 weeks, 0, 1000, 2000 and 4000 ppm (m: ca. 175, 400, 800 mg/kg bw/d; f: ca. 225, 425, and 850 mg/kg/d bw, LOAEL: 175 mg/kg bw/d (m); 225 mg/kg bw/d (f) (reduced bw) (NTP, 1998) Oral (gavage) study, mouse, 14 weeks, 0, 75, 150, and 300 mg/kg bw/d, NOAEL: 75 mg/kg bw/d (m); 150 mg/kg bw/d (f) (mortality, reduced bw) (NTP, 1998)	Feeding study, rat (SD), 13 weeks (non-GLP), 0, 0.02, 0.1, 0.2 % (equivalent to 0, 25, 125, and 250 mg/kg bw/d). NOAEL: 125 mg/kg bw/d (m: significant reduction in bw gain and absolute testes weight (Tarka, 1982)	-
Reproductive toxicity / Developmental Toxicity / Teratogenicity	Rat (CD), oral (diet), 0, 1500, 3000, 4000 ppm = 0, 124, 218, 259 mg/kg bw/d, gestational days (GD) 6-15, NOAEL maternal toxicity 124 mg/kg bw/d, NOAEL fetotoxicity 124 mg/kg bw/d, NOAEL teratogenicity 259 mg/kg bw/d (OECD, 2001) Mouse (CD-1), oral (drinking water), 0, 750, 1500 or 2000 ppm (= 282, 372, 396 mg/kg bw/d) on GD 6-15, NOAEL maternal toxicity: 750 ppm (282 mg/kg bw/d), NOAEL fetotoxicity: 750 ppm (282 mg/kg bw/d), NOAEL teratogenicity: 2000 ppm (396 mg/kg bw/d) (OECD, 2001) Rabbit, intravenous (IV, automatic infusion pump), 0 15, 30 and 60 mg/kg bw/d (maternal C_{max}: 30, 56 and 106 µg/mL), GD 6-18, NOAEL maternal toxicity: 30 mg/kg bw/d, NOAEL fetotoxicity / teratogenicity: 30 mg/kg bw/d (Shibata *et al*, 2000)	Rat (Sprague-Dawley), 0, 250 and 500 mg/kg bw/d, oral (gavage), 2 and 4 weeks, NOAEL < 250 mg/kg bw/d (testicular toxicity) (Theocorp Holding Company, 2010) Rat (Sprague-Dawley), 0, 0.0625 or 0.135 % (equivalent to 53 or 99 mg/kg bw/d), oral (diet), GD 6-19, NOAEL maternal toxicity: 53 mg/kg bw/d, NOAEL fetotoxicity / teratogenicity: 99 mg/kg bw/d mg/kg bw/d (Theocorp Holding Company, 2010) Rabbit (New Zealand), 0, 25, 75, 125, 200 mg/kg bw/d, oral (gavage), GD 6-29, NOAEL maternal toxicity: 75 mg/kg bw/d, NOAEL fetotoxicity / teratogenicity: 25 mg/kg bw/d (Theocorp Holding Company, 2010) Rabbit (New Zealand), 0, 0.0625, 0.125, or 0.188% (approx. 0, 21, 41, or 63 mg/kg bw/d), oral (diet), GD 6-29, NOAEL maternal toxicity: 21 mg/kg bw/d, NOAEL fetotoxicity / teratogenicity: 21 mg/kg bw/d (Theocorp Holding Company, 2010)	Mice (C57BL/6J), 0, 175 or 300 mg/kg bw (dissolved in deionised water), IP, GD 11-12, NOAEL maternal toxicity: 175 mg/kg bw, NOAEL fetotoxicity / teratogenicity: <175 mg/kg bw (York *et al*, 1986)

Conclusion:

Following repeated oral administration in rats, theophylline caused nephropathy in male rats and a dose- dependent periarteritis in all treated rats. Periarteritis was not observed in mice. The particular sensitivity of rats is most probably due to their anatomical situation as compared to mice and men. Since the periarteritis is considered a rat-specific response to vasodilators (the pathogenesis of theophylline-induced vascular lesions may be a consequence of hemodynamic changes induced in the vascular wall), it is of little, if any, relevance to humans (Nyska *et al*, 1998). Furthermore, this effect has not been associated with theophylline treatment in humans (OECD, 2001). Theophylline was not shown to be teratogenic in CD-1 mice at oral doses up to 396 mg/kg bw/d or in CD-1 rats at oral doses up to 259 mg/kg bw/d. Intravenous theophylline was fetotoxic and teratogenic in rabbits at maternal toxic doses exceeding the effective therapeutic range (60 mg/kg bw/d IV). At an

oral dose of 218 and 396 mg/kg, fetotoxicity was observed in rats and mice, respectively, in the presence of maternal toxicity.

Following repeated oral administration in rats, theobromine caused testicular toxicity with a NOAEL of 125 mg/kg bw/d. Theobromine was not shown to be fetotoxic or teratogenic in rats up to oral doses of 99 mg/kg bw/d. The NOAEL in rabbits for developmental toxicity was defined at 20-25 mg/kg bw.

Paraxanthine has been reported to induce malformations in mice following IP treatment at 175 and 300 mg/kg bw on day 11 and 12 of gestation. No NOAEL was established.

Conclusion – For the further evaluation, the lowest NOAEL value from the developmental toxicity study with rabbits of 30 mg/kg bw/d for theophylline will be taken. This NOAEL is based on the fetotoxicity/teratogenicity together with maternal toxicity. For this study, blood concentration of theophylline is known.

5.2.3. Supportive data (ToxCast)

The identification of structurally similar analogues was supported by a structural similarity search using the ToxCast database with a similarity cutoff of greater than 70 (>2700 chemicals were pulled back from GRASP, 10 of the 34 ToxCast chemicals had hits in ToxCast). The target families with positive hits for caffeine were the cell cycle, nuclear receptor, DNA binding, GPRC and esterase groups. Two analogues, theophylline and theobromine showed very high similarity among the chemicals tested by ToxCast and the most common hits to caffeine which supports the analogue selection (Table 9).

Table 9. Identification of structural analogues of caffeine with ToxCast data

Name	CAS	Similarity Cutoff (Isentris)	Result Count (All)	Result Count (Hit)	Positive hits details
Caffeine	58-08-2	100	663	21	2 background measurement, 5 nuclear receptor, 6 DNA binding, 2 cell cycle, 3 gpcr, 2 esterase, 1 signalling
Theophylline	58-55-9	>80	770	13	2 background measurement, 2 nuclear receptor, 4 DNA binding, 1 cell cycle, 3 gpcr, 1 esterase
Theobromine	83-67-0	>90	716	12	1 nuclear receptor, 1 DNA binding, 2 gpcr, 2 oxidoreductase, 2 cyp, 1 protease, 1 ion channel, 1 transporter, 1 misc
Etofylline	519-37-9	>90	109	2	1 background measurement, 1 DNA binding
Theofibrate	54504-70-0	>80	64	2	1 nuclear receptor, 1 cell cycle
Dimenhydrinate	523-87-5	>70	113	5	2 background measurement, 2 DNA binding, 1 nuclear receptor
Valacyclovir hydrochloride	124832-27-5	>70	109	1	Background measurement
8-Bromotheophylline	10381-75-6	>80	109	1	Nuclear receptor
Enprofylline	41078-02-8	>80	64	1	Cell cycle
Isbufylline	90162-60-0	>90	109	1	Background measurement

Caffeine had 21 hits in the ToxCast assays. Two analogues, theophylline and theobromine showed the most common hits to caffeine which supports the analogue selection

5.2.4. Estimation of the internal concentration of caffeine after dermal and oral exposure by physiologically-based biokinetic (PBBK) modelling

Physiologically-based Biokinetic (PBBK) Model

PBBK models are mathematical models used to quantify the absorption, distribution, metabolism and excretion of a chemical inside the body following exposure. They are constructed as an interconnected system of compartments representing various tissues described by mass balance differential equations that are solved to predict the amount of chemical in each compartment over time (Gerlowski and Jain 1983). The physiological basis of this modelling approach allows internal concentrations resulting from external exposures to be predicted, allowing comparisons including across species and exposure routes.

A number of recent reviews of PBBK modelling in environmental risk assessment are available (Clewell and Clewell 2008, Campbell *et al.* 2012, Clewell *et al.* 2014). A typical equation for a perfusion-limited tissue describes the mass balance for the uptake and clearance of the chemical in the tissue, in this case the liver:

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

This equation can be interpreted as: The rate of change in the mass of the chemical (A_{Liver}) in the liver is equal to the liver blood flow (Q_L) multiplied by the difference between the concentrations in the blood entering and leaving the liver ($C_{Arterial} - C_{Venous}$), minus the metabolic clearance of the chemical in the liver (Cl_{Liver}).

These models typically rely on three types of parameters; physiological (e.g. tissue volumes, blood flows), physicochemical (e.g. octanol:water portioning, vapour pressure, water solubility), and biochemical (e.g. absorption rates, metabolism, clearances). The particular parameters needed depend on factors such as the chemical properties and the purpose of the model. Various guidance documents for the application, use, and reporting of PBBK models have been published (WHO, 2010; U.S. EPA, 2006; U.S. FDA, 2018).

The physiological structure of PBBK models provides a particularly useful framework for conducting cross species extrapolations (Clewell and Andersen 1985). The necessary physiological parameters (tissue weights, blood flows, ventilation rate) for a number of mammals (mouse, rat, dog and human) are available in the literature (Brown *et al.* 1997), and parameters for other species can be estimated allometrically (Lindstedt and Schaeffer 2001). Tissue:blood partition coefficients for a chemical can be estimated using quantitative structure-property relationships (Peyret *et al.* 2010), while the clearance of the chemical in different species can be determined by *in vitro* studies with hepatocytes or cellular fractions and incorporated into the PBBK model using *in vitro* to *in vivo* extrapolation (Yoon *et al.* 2012).

The application of PBBK models to support interspecies extrapolation depends on the concept of target tissue exposure equivalence; that is, in the absence of pharmacodynamic (susceptibility) differences, the toxicity of a chemical in different species is expected to be associated with similar concentrations of the chemical (or its toxic metabolite) in the tissue where the toxicity is observed (Clewell *et al.* 2002). In cases of general systemic toxicity, or where the target tissue has not been identified, the concentration in the blood can be used to represent the target tissue exposure. While acute effects may depend on the maximum concentration achieved in the tissue, longer-term toxicity is generally associated with the average concentration over time, which can be calculated as the area under the curve (AUC) divided by the duration of the exposure. The toxic mode of action determines whether the

concentration of interest is that of the parent chemical, a stable metabolite, or a reactive metabolite (Clewell 2005).

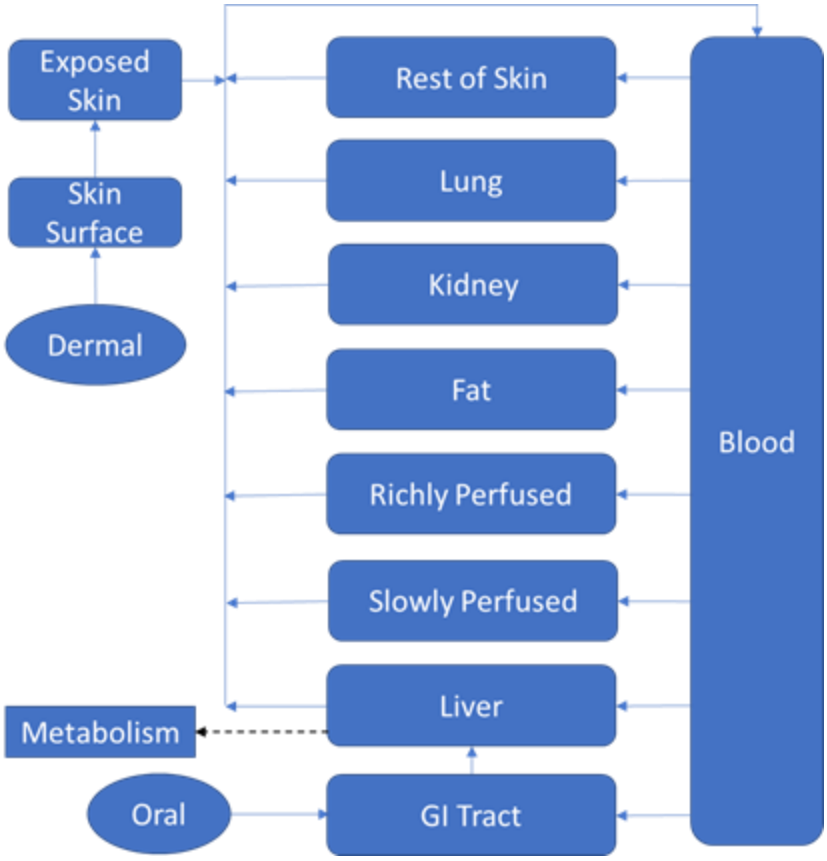
To apply a PBBK model for interspecies extrapolation, the model is first used to simulate the exposure of interest (dose, route, and duration) in the experimental species, and the internal dose metric (peak or average concentration) is calculated. The parameters in the PBBK model are then changed to those for the target species of concern and the dose is adjusted until the same internal dose metric is achieved. The dose that produces the same internal dose metric is then considered the kinetically equivalent dose (Clewell *et al.* 2002).

Model Structure

A physiologically-based biokinetic (PBBK) model was developed in Berkeley Madonna software (version 8.3.18; University of California, Berkeley, CA; www.berkeleymadonna.com). It is structured similarly to that reported by Gajewska (2015), with perfusion-limited compartments for skin, liver, fat, lung, kidney, blood, and lumped compartments for the remaining richly and slowly perfused tissues. The major differences with the Gajewska model are that a simplified model was developed (i.e., single GI and skin compartments), and the mass balance equations were corrected or rewritten. The simplified model results are very similar to the more complicated, multi-compartment models reported by Gajewska, so additional complexity was not added. All tissue volumes and blood flows have been updated to match average human values reported in Brown *et al.* (1997, Tables 21 and 23). There is little change in C_{max} (10.42 to 10.38 mg/L) or AUC (174.8 to 175.0 mg-h/L).

Consistent with common practice, tissue:venous equilibration is assumed, and the tissues are assumed to be well-mixed reservoirs. Exposure is characterised in exposed skin (dermal) and gastrointestinal (GI, oral) compartments, and metabolism is described as a first-order clearance process in the liver. The model structure is shown schematically in Figure 3.

Figure 3. PBBK model schematic for caffeine showing the representation of the main organs considered with various sub-compartments in the skin and GI tract for oral and dermal exposure



Tissue blood flows and volumes are set to values in Brown *et al.* (1997, Tables 21 and 23), except for skin. Skin volume was calculated as the product of surface area and average skin thickness. Skin thickness was taken from Brown *et al.* (1997) and skin surface area was calculated using the following allometric relationship from Livingston and Lee (2001):

$$SA = 0.1173 * BW^{0.6466} \ (m^2).$$

Blood flow to the exposed skin was calculated as the total skin blood flow adjusted by the ratio of volume exposed skin to volume of total skin. Exposed skin volume and blood flow was subtracted from the total skin to derive the parameters for the unexposed skin compartment. A list of physiological parameter values is shown in Table 10.

Table 10. Human PBBK physiological parameter values for caffeine. Source for values is Brown *et al.* (1997) unless otherwise specified.

Parameter	Units		Symbol	Value
Body mass [d]		kg	BW	-
Skin thickness		cm	Depth	0.1
Blood Flows (Fraction of Cardiac Output) [a]				
Cardiac Output		L/h/kg$^{0.75}$	QCc	15
Fat		1	QFc	0.052
Kidney		1	QKc	0.75
Liver		1	QLc	0.227
Skin		1	QSkc	0.058
Lung		1	QLuc	0.025
Volumes (fraction of BW) [b]				
Fat		1	VFc	0.21
Kidney		1	VKc	0.004
Liver		1	VLc	0.026
Lung		1	VLuc	0.008
Blood [c]		1	VAc, VVc	0.079
Hepatocellularity [e]	millions of hepatocytes per gram liver		hpgl	99

a Richly perfused blood flow = 70% of QC minus liver, kidney, and lung volumes.
Slowly perfused blood flow = 30% of QC minus fat, and skin volumes.
b Richly perfused tissue volume = 70% of BW minus liver, kidney, and lung volumes.
Slowly perfused tissue volume = 83.6% of BW minus fat, blood, and skin volumes.
c Blood is divided into 3/4 arterial and 1/4 venous.
d Simulation specific
e Barter *et al.* (2007)

Tissue:blood partition coefficients (PC) were estimated by Gajewska using the algorithm of Schmitt (2008), except for skin. Skin:blood PC was estimated as a weighted average of liver and fat PCs (i.e. 0.7*PL + 0.3*PF).

Oral exposure is modelled using a 1-compartment, first-order absorption model, with the gastrointestinal (GI) tract acting as a reservoir for the oral dose. All oral doses are simulated as a bolus doses (i.e. all chemical ingested at once per dosing event). The blood flow from the GI enters the liver via the portal vein.

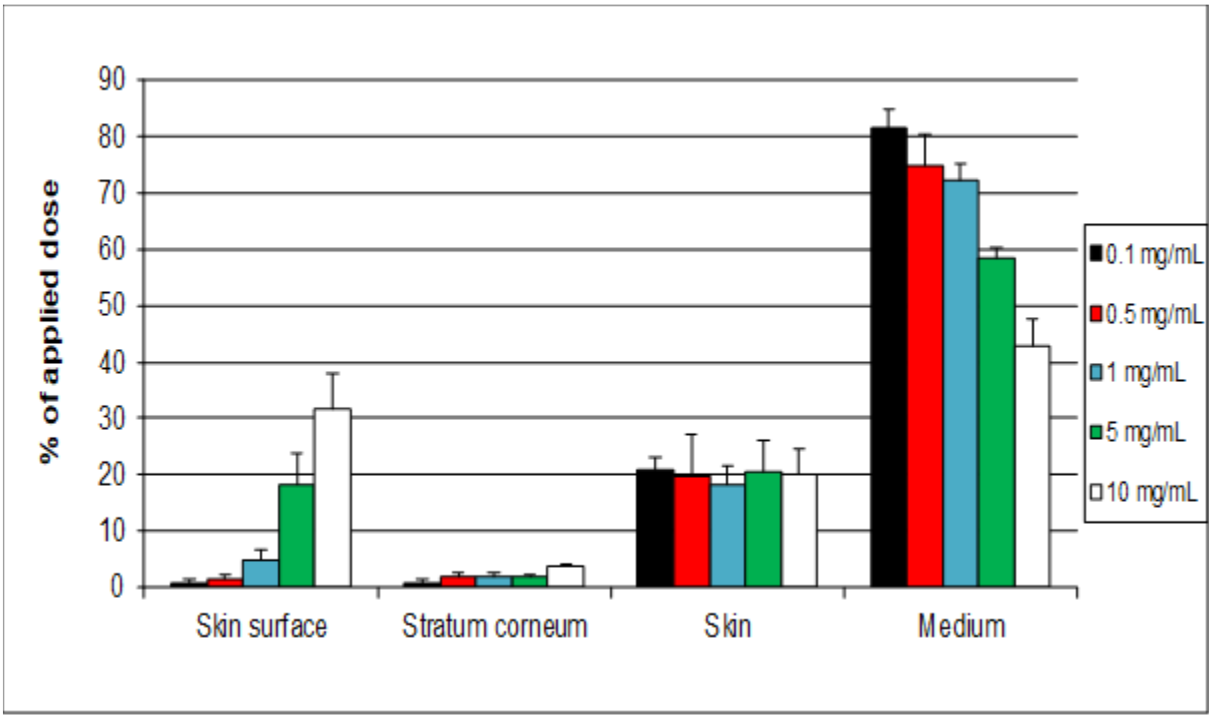
Dermal exposure is modelled using a skin surface compartment to house the applied dose in terms of the volume and surface area, and a single skin compartment. Transfer to the systemic circulation occurs in the skin compartment. The skin is separated into exposed and unexposed compartments. Dermal absorption is driven by a permeability coefficient for uptake from the surface into the skin (units of cm/hr), the exposure area and volume of application, the amount of chemical applied, duration of application, and the fraction absorbed. Transfer from the skin to the blood is modelled assuming a well- mixed, blood flow-limited exposed skin compartment and a skin:blood partition coefficient.

To simulate the experimental dermal exposure data from Otberg *et al.* (2008), applied volume and concentration was calculated from details reported in the exposure methodology. For the consumer whole-body cosmetic exposure scenario, the applied volume was assumed to be 4 mL (Troutman *et al.*, 2015), the surface area exposed was taken to be the whole skin surface area, and the concentration was assumed to be 2% in lotion. The fraction available for dermal absorption was extrapolated based on a 2% concentration and *in vitro* dermal penetration data collected by Cosmetics Europe (2017) (Figure 4). Specifically, assuming a lotion density of 1 g/mL, 2% corresponds to 20 mg/mL,

which is higher than the greatest concentration tested of 10 mg/mL. Inspection of the graph shows decreasing penetration into the medium as concentration increases. As a conservative value, the 10 mg/mL absorption, which is close to 50%, was used to estimate availability at the estimated simulated concentration of 20 mg/mL.

Figure 4. Dermal penetration after 24h of caffeine applied in a range of concentrations (0.1 – 10 mg/mL) to viable human skin (Hewitt *et al*, 2019).

The skin penetration rates (as % of applied dose) are shown for the skin surface, stratum corneum, skin (epidermis and dermis) and the medium (receptor fluid)



The intrinsic hepatic clearance and oral absorption rate constant were fitted to experimental data collected in controlled oral and dermal human exposures (Denaro *et al.*, 1991; Lelo *et al.*, 1986; Otberg *et al.*, 2008). The parameters were fit simultaneously to the individual data sets, and the average fitted value was used for all subsequent simulations. Chemical-specific parameters for caffeine are shown in Table 11.

Table 11. Chemical-specific PBBK parameters for caffeine

Parameter	Units	Symbol	Value	Source
Molecular weight	g/mol	MWC	194.2	PubChem
Partition Coefficients				
Fat	1	PF	0.68	Schmitt (2008)
Kidney	1	PK	3.76	Schmitt (2008)
Liver	1	PL	4.25	Schmitt (2008)
Lung	1	PLu	1.23	Schmitt (2008)
Rich	1	PR	2.4	Schmitt (2008)
Slow	1	PS	0.995	Schmitt (2008)
Blood and Plasma				
Fraction unbound in blood	%	fub	96	Lave *et al.* (1997)
Fraction unbound in plasma	%	fup	68	Lelo *et al.* (1986)
Blood:plasma ratio	%	RBP	71	fup/fub
Oral absorption				
GI -> liver	1/h	Ka	1.6	fit
Dermal absorption				
Fraction available	%	FracAvail	50	Géniès *et al*, Hewitt *et al.(2019)*
Permeability	cm/h	Kp	4.10E-04	Doucet *et al.* (1998)
Metabolism				
Hepatocyte clearance	uL/min/million cells	hep_$Cl_{int_in\ vitro}$	0.68	fit

The liver clearance was calculated by scaling the *in vitro* intrinsic clearance value to the whole body. The liver clearance rate is calculated as the uL/min/million cells, times hepatocellulrity, times the volume of the liver:

$$\mathrm{CL_{int}(L/h)} = \mathrm{CL_{int\_in\ vitro}} * \mathrm{hpgl} * 10^{-6}(\mathrm{L/\mu L}) * 60\ (\mathrm{min/h}) * 10^{3}(\mathrm{g/kg}) * \mathrm{VLc} * \mathrm{BW(kg)}$$

where $CL_{int\_in\ vitro}$ is the uL/min/million cells cleared *in vitro*, hpgl is the number of million hepatocytes per gram of liver, and VLc*BW is the volume of the liver.

A local sensitivity analysis for model parameters was conducted using the built-in tool in the Berkeley Madonna software. The sensitivity coefficients were then normalised by the output and input parameter values according to the following equation:

$$\text{Normalised Sensitivity Coefficient} = \frac{\Delta Y/Y}{\Delta X/X}$$

where Y is the output (i.e., Cmax or AUC), X is the input parameter (e.g., K_a, K_p), ΔX is the change in the parameter value, and ΔY is the resulting change in the output value. Normalisation of the sensitivity coefficients is necessary to make comparisons across parameters of different scales (Clewell *et al.* 1994).

Results

A qualitative evaluation of the agreement between experimental plasma concentration and simulations was conducted through visual inspection of the time-course concentration curves (U.S. EPA, 2006). Good agreement, with model predictions generally within a factor of two of the data (WHO 2010) was obtained for both the oral and dermal route (Figure 5, Figure 6 and Figure 7). Only the hepatocyte clearance (CL_{int}) and oral absorption coefficient (K_a) are fitted to experimental data, while all remaining parameters were obtained from the literature or data collected by Cosmetics Europe (2017). The values of CL_{int} obtained from fitting to the data of Denaro *et al.* (1991) (oral) and Otberg *et al.* (2008)

(dermal) are very similar (0.7 vs 0.5 ul/min/million hepatocytes). A value of 1.4 µl/min/million hepatocytes was reported by Lelo *et al* (1986), which is also close to the fitted values.

Figure 5. Caffeine PBBK simulations compared to measured repeated oral exposure data in human volunteers (Denaro *et al.*, 1991).

Caffeine plasma concentrations (black stars, green squares) were determined following ingestion of 6 cups of caffeinated coffee (0.7 and 2 mg/kg caffeine) over 5 days. Since the body weight (BW) was not reported, a default value of 70 kg was used.

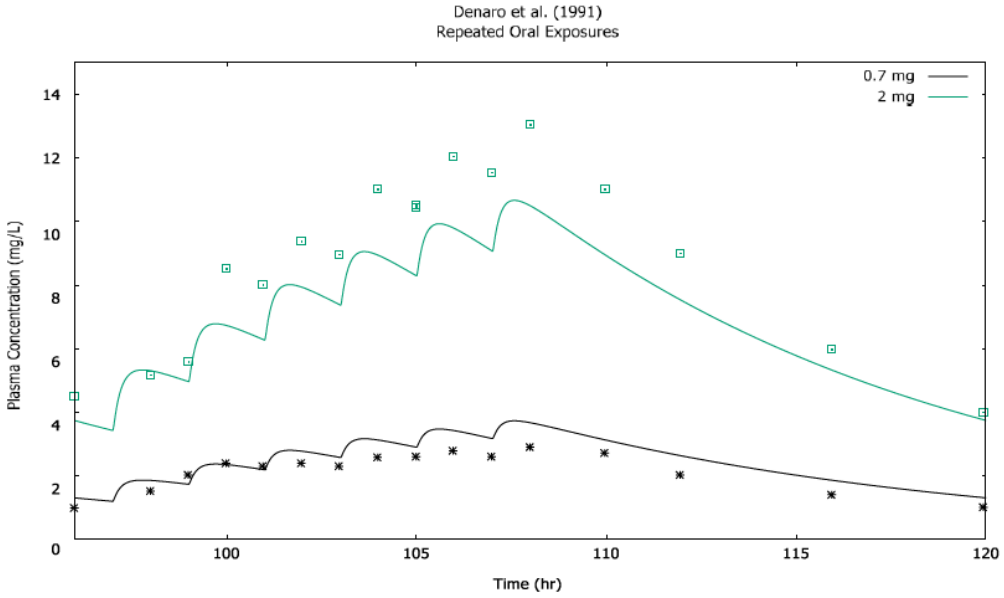


Figure 6. Caffeine PBBK simulations compared to measured oral exposure data in human volunteers (Lelo *et al.*, 1986).

Caffeine plasma concentrations (black crosses) were determined following ingestion of 3.25 mg/kg bw caffeine in a gelatine capsule. BW = 83 kg.

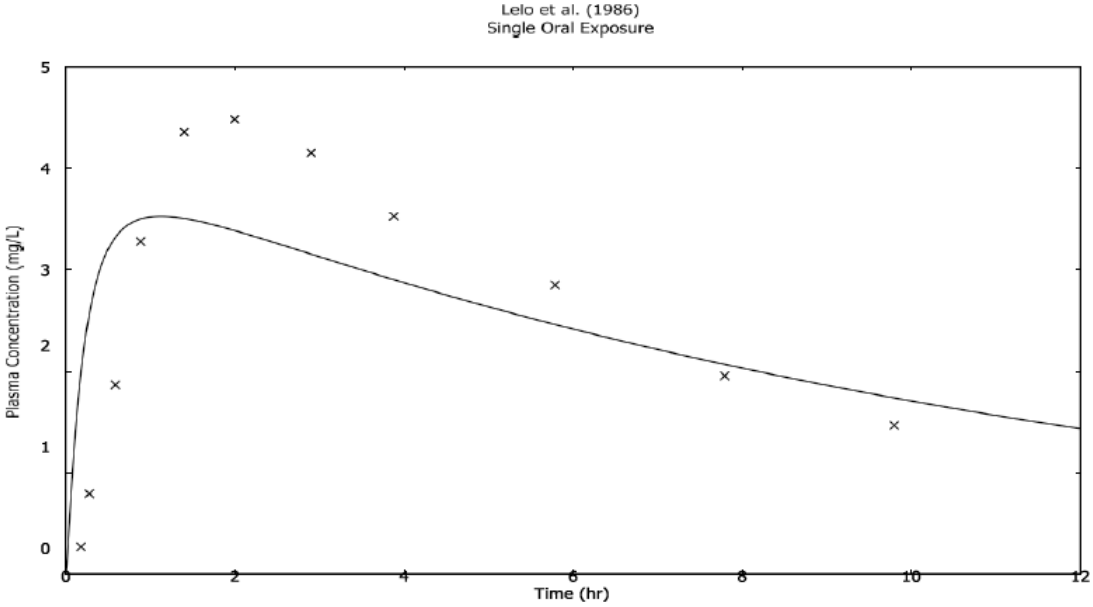
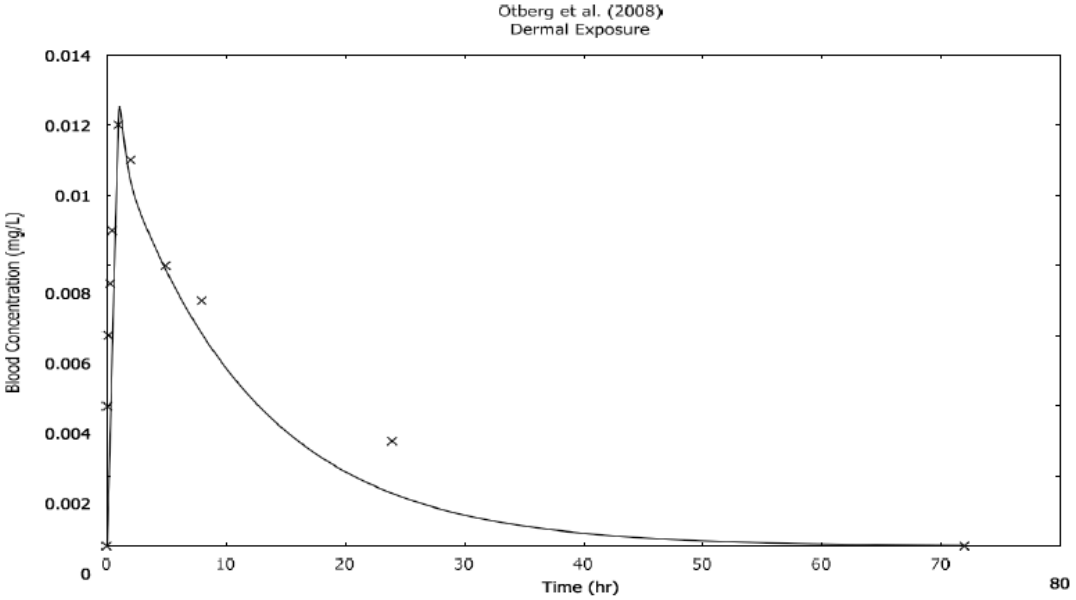


Figure 7. Caffeine PBBK simulations compared to measured dermal exposure data in human volunteers (Otberg *et al.*, 2008).

Caffeine plasma concentrations (black crosses) were determined following dermal application of 50 mg (0.05 mL) of an ethanol/propylene glycol formulation containing 2.5% caffeine to 25 cm^2 of the chest of 6 male volunteers (1.25 mg caffeine). Since the body weight (BW) was not reported, a default value of 70 kg was used.



A simulation of the worst-case exposure estimated for caffeine via oral and dermal routes was conducted (Figure 8). Upper bound oral (13.1 mg/kg/d) and dermal (4.6 mg/kg/d) exposure estimates (see section 2.4) were determined and used as input to the model. Twice daily exposures were simulated, 12 hours apart, as bolus ingestions or applications. A 4 mL volume was used to simulate whole body exposure to lotion, as reported by Troutman *et al.* (2015). Caffeine concentration in the dermal formulation was assumed to be 2%, which corresponds to approximately 20 mg/mL caffeine. Using the dermal penetration trend shown in Figure 4, an average value of 50% dermal absorption was assumed for caffeine. A default body weight of 60 kg was used as recommended in the SCCS NoG (2018).

The simulated internal dose metrics (C_{max} and AUC) are shown in Table 12. The dosing simulation was run for 3 (simulated) days to achieve steady periodicity, as shown in Figure 8. The maximum concentration in blood (C_{max}) was 10 mg/L, the daily area under the concentration curve (AUC, calculated over the time period from 48 hours to 72 hours) was 170 mg*h/L, and the average daily concentration (C_{avg}) was 7.3 mg/L.

Table 12. Summary of the worst-case cosmetic exposure scenario.

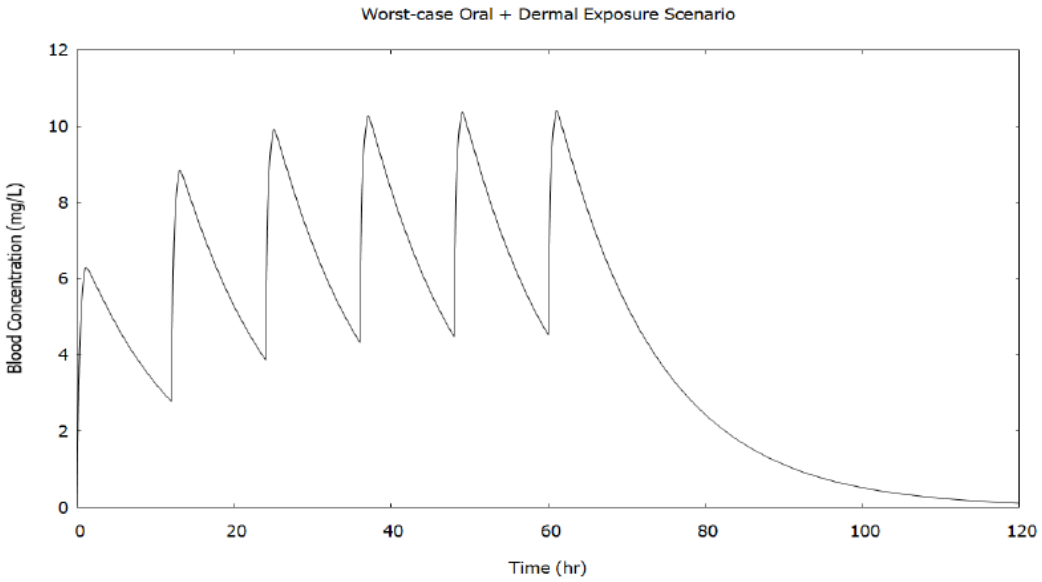
Daily oral and dermal doses are split into 2 equal doses, administered at 12-h intervals.

Exposure	
Single Oral Dose (mg/kg)	6.55
Single Dermal Dose (mg/kg/d)	2.4
Dermal Exposure Area cm^2	16560[a]
Doses/day	2
Daily Oral Dose (mg/kg/d)	13.1
Daily Dermal Dose (mg/d)	4.6
Internal Dose Metrics	
Cmax (mg/L)	10
AUC (mg-h/L)	170
Cavg (mg/L)	7.3

[a] Whole body surface area for a 60 kg human calculated using the equation of Livingston and Lee (2001)

Figure 8. PBBK simulation of internal plasma concentration of caffeine following estimated oral and whole body dermal exposure.

Oral exposure = 13.1 mg/kg/d, whole body dermal exposure = 4.6 mg/kg/d in 4 mL of product, twice daily exposure. BW = 60 kg. C_{max} = 10 mg/L, AUC = 170 mg*h/L, C_{avg}= AUC/24 = 7.3 mg/L.



The results of the sensitivity analysis for C_{max} and AUC are shown graphically in Figure 9 and Figure 10. The most influential parameters and their coefficients are tabulated in Table 13. A positive sensitivity coefficient indicates that as the parameter increases the dosimeter increases, while a negative coefficient indicates a decrease in output as the parameter increases. The model is generally positively affected by increases in parameters driving absorption, and negatively by increases in parameters driving clearance.

Figure 9. Normalised sensitivity coefficients for C_{max} (T_{max} = 61 h).

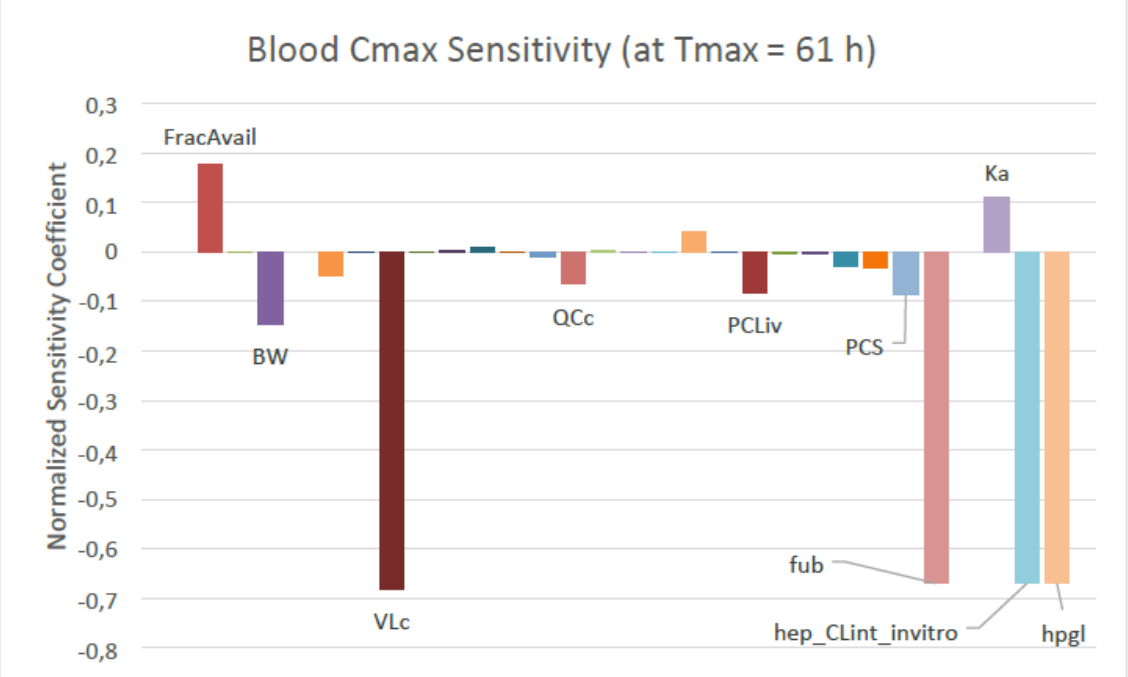


Figure 10. Normalised sensitivity coefficients for AUC

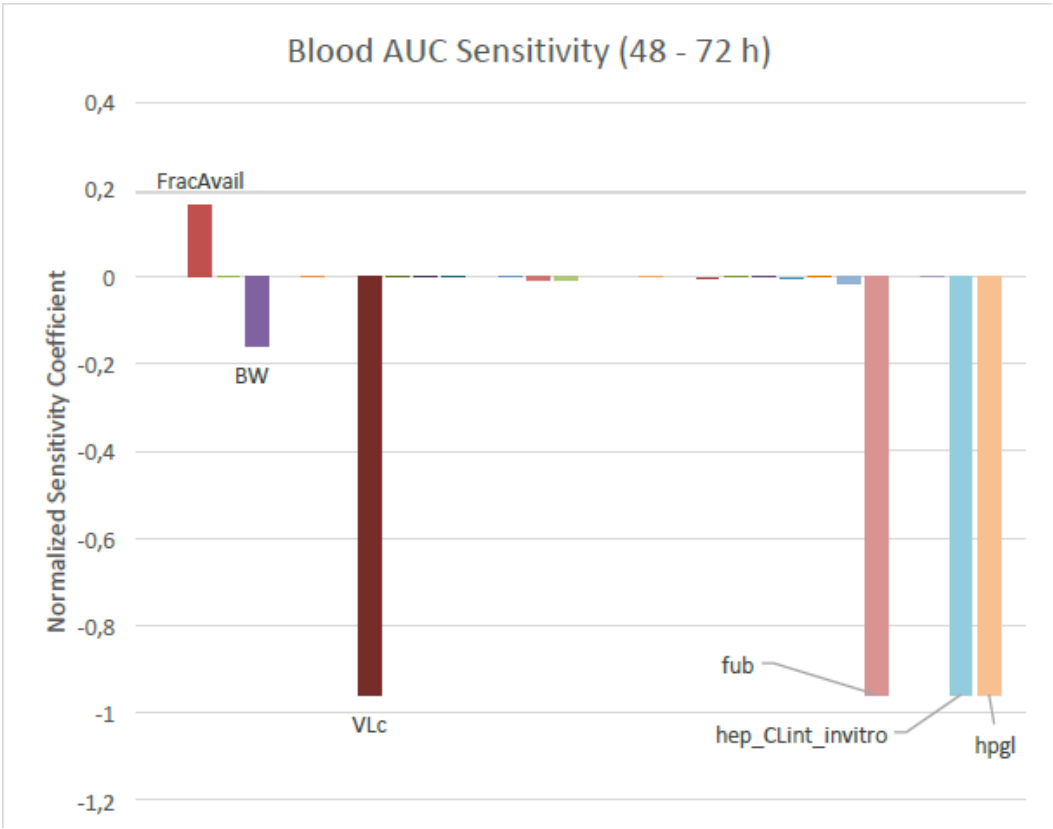


Table 13. Sensitivity Coefficients for C_{max} and AUC. Parameters with coefficients greater than 0.05 or less than -0.05 are shown.

Parameter	Symbol	Cmax Sensitivity	AUC Sensitivity
Fraction Absorbed	FracAvail	0.18	0.16
Body Weight	BW	-0.15	-0.16
Liver Volume	VLC	-0.68	-0.96
Cardiac Output	QCC	-0.06	-
Liver:Blood Partition	PCLiv	-0.08	-
Slowly-perfused:Blood Partition	PCS	-0.09	-
Oral Absorption Rate	Ka	0.11	-
Fraction Unbound in Blood	fub	-0.67	-0.96
Hepatic Intrinsic Clearance	hep_$CL_{int_invitro}$	-0.67	-0.96
Hepatocellularity	hpgl	-0.67	-0.96

*Conclusions: After oral caffeine exposure (13.1mg) and dermal exposure (4.6mg) the maximum concentration in blood (C_{max}) was 10 mg/L, the daily area under the concentration curve (AUC, calculated over the time period from 48 hours to 72 hours) was 170 mg*h/L, and the average daily concentration (C_{avg}) was 7.3 mg/L.*

Discussion of Results and Uncertainties of PBBK model

A PBBK model with relatively simple oral and dermal absorption models was developed. The model simulations agreed reasonably well across multiple oral and dermal exposure data sets using a common set of parameter values. The model did tend to underpredict the plasma concentration at the higher oral doses, though it was consistently within a factor of 2 from the data. This is likely due to saturation of oxidative metabolism, and there are also potential vehicle effects that are not being captured by the model (Lelo *et al.*,1986) used gelatine capsules, while Denaro *et al.* (1991*)* put anhydrous caffeine into decaffeinated coffee). A more complex description of the oral absorption model could potentially improve the accuracy of the simulations at high doses, though more data or assumptions would be required to parameterise the model, bringing with them uncertainties of their own.

The fraction available for dermal absorption was set to 50% based on extrapolation of the trend found using *in vitro* testing. It is noted, however, that the fraction absorbed has been observed as high as 67%, which would lead to higher predictions of C_{max} and AUC. To estimate the impact of this assumption, the model was run again with 67% available for absorption, and the results are shown in Table 14. There is a 10% increase in C_{max}, and approximately a 6% increase in AUC.

Table 14. Comparison of internal dose metrics with 67% dermal absorption vs 50%.

Fraction Absorbed	50%	67%
C_{max} (mg/L)	10	11
AUC (mg-h/L)	170	180

The sensitivity analysis identified 6 parameters to which the AUC and C_{max} were sensitive, and an additional 4 to which C_{max} only was sensitive. The parameters driving absorption (FracAvail and Ka) had positive coefficients (i.e., increased parameter results in increased

C_{max}), while parameters increasing clearance (e.g., liver parameters and intrinsic clearance) had negative coefficients.

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

5.3. Justification

The caffeine analogues theophylline, theobromine and paraxanthine form a robust category for the read across because:

- Based on the Tanimoto score they show a very high degree of structural similarity ≥ 0.9,
- They have very similar measured physico-chemical properties
- The metabolism data confirm that caffeine is metabolised *in vivo* to those 3 compounds with paraxanthine being the major metabolite (80%)
- The *in silico* profiling supports the hypothesis that they have similar characteristics with respect to bioavailability, solubility, logP
- The QSAR toolbox, VEGA and endocrine disruptome platforms predict the same toxicological characteristics for the target and the source compounds
- They have a similar mode of action, i.e. adenosine receptor inhibition, with K_i factors that increase from theobromine < caffeine < paraxanthine < theophylline

The additional information gathered with new approach methodologies (NAM) was provided by *in silico* safety alert tools, which did not identify substantial toxicity alerts for caffeine or its three major metabolites, *in vitro* metabolism data and by PBBK modelling estimating the systemic exposure following potential aggregate exposure to caffeine-containing products by both the dermal and oral route. The *in silico* and *in vitro* data supported the hypothesis that theophylline, theobromine and paraxanthine are suitable analogues and theophylline was the closest caffeine analogue.

Several repeated dose toxicity and/or reproductive toxicity animal data could be gathered for all three analogues theophylline, theobromine and paraxanthine. Most of the studies were judged of sufficient quality and have been summarised in this document (Table 8, Annex 3).

The common MoA between caffeine and the closest structural analogues was considered to be adenosine receptor antagonism with theophylline having the highest relative potency. In pregnant women, they can cause increase in maternal levels of 3'5'-cyclic adenosine monophosphate and epinephrine, which could lead to uteroplacental vasoconstriction and decreased intervillous placental blood flow, which could restrict foetal growth. Therefore, the NOAEL of the developmental toxicity study in rabbits after IV injection was selected

as point of departure (PoD) for the safety assessment being the lowest NOAEL among the reported studies taking into account study duration and exposure route. The adverse effects in the developmental toxicity study in rabbits occurred in parallel with maternal toxicity and the mechanism is not specified. Despite the fact that it is not possible to conclude that the MoA causing the developmental toxicity in rabbits is the same MoA as the pharmacological effects of caffeine and its analogues in human, this choice of PoD is considered to be very conservative and consequently protective for human health.

6. Strategy for and integrated conclusion of data gap filling

6.1. Uncertainty

Factor	Uncertainty (low, medium, high)	Comment
Hypothesis used for the read across	Low	Three analogues (theophylline, theobromine, paraxanthine) have been found on the basis of comparable structural, physico-chemical and molecular properties. In addition, based on available data they are considered to have a common mode of action (MOA) with different potencies
Structural similarity	Low	High level of structural similarity (all are methyl xanthines). Tanimoto score >0,9 based on CE-ToxGPS software.
Similarity of physico- chemical properties	Low	The 4 chemicals (caffeine, theophylline, theobromine, paraxanthine) have similar physico-chemical properties (similar MW and logP range, similar negligible volatility and similar solubility classes), all are predicted to be bioavailable and to have a good skin penetration
Toxicokinetic similarity	Low	Following oral intake, caffeine, theophylline, theobromine and paraxanthine are readily and completely absorbed. The biotransformation is catalysed by the cytochrome P-450 monooxy-genase system and leads to the formation of demethylated compounds or, through oxidation, to the urate and/or hydration to the diaminouracil.
Mode of action (MoA)	Low	The common pivotal MoA between caffeine and the source chemicals was considered to be antagonism of the A1-adenosine receptor based on pharmacological investigations. The differences in the relative potency were taken into account.
Similarity of other supportive data (ToxCast data)	Low	Theophylline and theobromine have been identified as the closest structural analogues among all the chemicals in ToxCast. The categories of positive hits are similar to those of caffeine.
Number of analogues used for the read across	Low	The number of analogues used for the read- across was low, but considered sufficient. Valid *in vivo* data are available for theophylline that was considered the most similar analogue based on ToxCast and Tanimoto coefficient
Quality of the endpoint data used for the read across	Low	Several supporting *in vivo* studies with Klimisch score 1 & 2 available for all three analogues. The rabbit prenatal developmental toxicity study with theophylline was used since it was an intravenous study (100% bioavailability) with data on internal plasma levels
Similarity of the endpoint data (among source chemicals)	Medium	Differences were most probably due to different pharmacological potencies
Concordance and weight of evidence of all data used for justifying the hypothesis	Low	Combination of similarity in structures, ADME, biological and kinetic properties (see Annex 2: QSAR Toolbox v 4.3)
Overall level of confidence in the read- across approach	Low	Available *in vivo* data on caffeine support the read - across approach. Moreover, internal plasma levels derived from a reliable (Klimisch score 2) *in vivo* study with theophylline, the most potent analogue, were used as point of departure (NOAEL). The final Margin of Internal Exposure (MoIE) calculation was considered sufficiently conservative since it was based on estimated internal exposure with a PBBK model, rather than on external doses with inherent uncertainty related to route-to-route extrapolation and species/strain differences

The level of uncertainty was considered low since based on the considered common MoA between caffeine and the closest structural analogues, internal plasma levels derived from an acceptable *in vivo* animal study with theophylline were used as point of departure (NOAEL). The final Margin of Internal Exposure (MoIE) calculation was considered sufficiently conservative since it was based on estimated internal exposure with a PBBK model, rather than on external doses with inherent uncertainty related to route-to-route extrapolation and species/strain differences. Thus, calculation of internal exposures with a PBBK model can be used to replace the default uncertainty factor of 4 for interspecies differences in toxicokinetics.

6.2. Integrated conclusion

6.2.1. Determination of a Margin of Internal Exposure (MoIE)

In this case study, a Margin of Internal Exposure (MoIE) has been estimated for caffeine using a PBBK model to estimate blood concentrations following exposures to caffeine in experimental animals and humans. A MoIE differs from a traditional margin of exposure (MoE) in that it is calculated as the ratio of a measure of internal exposure, such as blood concentration or target-tissue dose, rather than a measure of external exposure concentration or ingested dose (Bessems *et al.* 2017). The ability to rely on a measure of internal rather than external exposure reduces the uncertainty in the risk assessment by incorporating chemical-specific information on the uptake, distribution, metabolism and excretion of the chemical in both the experimental animal and the human (Clewell *et al.* 2008). In particular, calculation of internal exposures with a PBBK model can be used to replace the default uncertainty factor of 4 for interspecies differences in toxicokinetic differences (IPCS 2005, WHO 2010). The U.S. EPA follows this practice in determining Reference Concentrations and Reference Doses (U.S. EPA 1994, 2011). Thus, a MoIE of 25 would be equivalent to the default MOE of 100, but with greater precision for the chemical of concern.

Determination of the internal exposure of theophylline, the most similar analogue to caffeine, at the most conservative PoD from animal studies

> NOAEL – rabbit, prenatal developmental toxicity study (intravenous injection): 30 mg/kg bw/d; C_{max} at this NOAEL was 56 µg/mL, equivalent to 311 µM (MW= 180.167 g/mol) (OECD, 2001; Shibata *et al*, 2000)

Determination of the internal exposure to caffeine following human aggregate dermal (cosmetic products) and oral (food/drink) caffeine exposure according to PBBK modelling:

> C_{max} (caffeine) 10 mg/L, equivalent to 51 µM (MW= 194.194 g/mol)

Calculation of the margin of internal exposure (MoIE), accounting for relative potency:

$$\mathbf{MoIE = C_{max}NOAEL\ animal\ study/(C_{max}human\ x\ RPF)}$$
$$\mathbf{= 311\ \mu M/51\ \mu M\ x\ 0.24\ = 25.4}$$

The objective of the case study on caffeine was to demonstrate how read-across can be applied to fill data gaps in an assessment of the potential risk for the consumer from exposure to caffeine by using *in vivo* data from structural analogues, while assuming that no *in vivo* repeated dose toxicity data were available for the target substance caffeine.

Furthermore, this case study demonstrates the value added by New Approach Methodologies (NAMs) such as *in silico* tools, *in vitro* ADME, PBBK modelling, read-across, thereby opening new perspectives to adapt the traditional safety assessment approach to a next generation risk assessment (NGRA) including an acceptable level of confidence.

6.2.2. Comparison of PoD derived in the CS with in vivo NOAEL Caffeine

In an attempt to prove that the outlined read-across approach works, the available *in vivo* data on caffeine were examined for comparison (Table 15, Annex 3). When comparing data from 3-month rat studies, i.e. lowest caffeine NOAEL from rat drinking water study 151 mg/kg bw vs lowest theophylline NOAEL from rat diet study <75 mg/kg bw (Table 8, Annex 3), caffeine was shown to be less toxic compared to its most potent analogue

theophylline. The NOAELs from the available developmental toxicity studies were not directly comparable due to different species and exposure routes (rabbit NOAEL 30 mg/kg bw IV for theophylline vs rat NOAEL 40 mg/kg bw oral gavage for caffeine), but are indicating a similar range of toxicity taking into account the different species sensitivities.

Table 15. Repeat-dose and reproductive toxicity data for caffeine (Klimisch score 2)

In vivo data	Caffeine
Repeat-dose toxicity	Oral (drinking water) study, rat, 90 days, 0, 188, 375, 750, 1500, and 3000 ppm (male: ca. 19.7, 42, 85.4, 151, 272 mg/kg bw/d; female: 23, 51, 104, 174, 287 mg/kg bw/d); NOAEL 1500 ppm (m: 151 mg/kg bw/d, f: 174mg/kg bw/d (reduced bw gain) (OECD, 2002)
Reproductive/ Developmental toxicity	Oral (gavage), rat, GD 1-19, LOAEL 40 mg/kg bw/d (maternal toxicity, significant decrease in bw gain), NOAEL 40 mg/kg bw/d (fetotoxicity, significant reduction in fetal bw gain), NOAEL 80 mg/kg bw/d (teratogenicity) (OECD, 2002)

However, the calculation of a MoIE on the basis of safety data with caffeine, i.e. direct comparison of the estimated internal human exposure following dermal cosmetic use and the PoD/NOAEL from a caffeine animal toxicity study, is not possible since in none of the available animal studies internal compound exposure was assessed. Due to insufficient data and resources, the generation of an animal PBK model for caffeine was not possible. Therefore, the read-across approach with the theophylline data was so valuable.

Derivation of MOE with animal data for caffeine

The derivation of a margin of external exposure with animal caffeine data goes beyond the scope of this case study which used read across and NAMs to demonstrate a safety assessment in the absence of repeated dose toxicity data for the target chemical. In addition, there is substantial data from human use of caffeine that have been reviewed elsewhere (EFSA 2015). The epidemiological data shows that intakes up to 400mg per day (about 5.7mg/kg bw per day) consumed throughout the day do not raise safety concerns for healthy adults in the general population (EFSA 2015). If we would use this exposure estimate together with the NOAEL derived from the prenatal developmental toxicity study in rats the margin of (external) exposure would be 7.

$$\text{MoE} = \text{PoD animal study}/\text{“Safe” Exposure} = 40\text{mg/kg bw} / 5.7\text{mg/kg bw} = 7$$

7. Postface

During the review process, the reviewers asked for a summary of the regulatory activities on the chemicals covered by the case study. Please note, that the case study has been developed separately (and in some cases prior) to those regulatory activities and that it is not intended per se to influence those activities. In addition, the goal of the present case study is not hazard identification but to present a risk assessment supported by read across and NAMs.

Regulatory activities on theophylline:

ECHA Registry of CLH intentions until outcome

https://echa.europa.eu/registry-of-clh-intentions-until-outcome/-/dislist/details/0b0236e180923e53

Canada's Cosmetics Hotlist

https://www.canada.ca/en/health-canada/services/consumer-product-safety/cosmetics/cosmetic-ingredient-hotlist-prohibited-restricted-ingredients/hotlist.html

8. References

Barter ZE, Bayliss MK, Beaune PH, Boobis AR, Carlile DJ, Edwards RJ, Houston JB, Lake BG, Lipscomb JC, Pelkonen OR, *et al.* (2007), Scaling factors for the extrapolation of *in vivo* metabolic drug clearance from *in vitro* data: Reaching a consensus on values of human microsomal protein and hepatocellularity per gram of liver. *Current Drug Metabolism*, Vol. 8, pp. 33–45.

Berggren E, White A, Ouedraogo G, Paini A, Richarz AN, Bois FY, Exner T, Leite S, van Grunsven LA, Worth A, Mahony C (2017), *Ab initio* chemical safety assessment: A workflow based on exposure considerations and non-animal methods. *Computational Toxicology*, Vol. 4, pp. 31-44

Berthou F, Ratanasavanh D, Riche C, Picart D, Voirin T, Guillouzo A. (1989), Comparison of caffeine metabolism by slices, microsomes and hepatocyte cultures from adult human liver. *Xenobiotica*. Vol. 19, No. 4, pp. 401-17.

Bessems JGM, 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*, Vol. 392, pp. 119-129

Blanchard J, Sawers SJ (1983). The absolute bioavailability of caffeine in man. *Eur J Clin Pharmacol*. Vol. 24, No. 1, pp. 93-98

Bonati M., Garattini S. (1984) Interspecies Comparison of Caffeine Disposition. *In: Dews P.B. (eds) Caffeine*. pp. 48-56. Springer, Berlin, Heidelberg. https://doi.org/10.1007/978-3-642-69823-1_3.

Bradford, J.C., Caldwell, J.A., Barbolt, T.A., Drobeck, H.P. (1983). Chronic administration of caffeine to two generations of rats. *Teratology*, Vol. 27, pp. 2-9.

Brown, R.P., Delp, M.D., Lindstedt, S.L., Rhomberg, L.R., Beliles, R.P (1997). Physiological parameter values for physiologically based pharmacokinetic models. *Toxicol. Ind. Health*, Vol. 13, No. 4, pp. 407–484.

Campbell J.L., Clewell R.A., Gentry P.R., Andersen M.E., Clewell H.J. (2012) Physiologically Based Pharmacokinetic/Toxicokinetic Modeling. In: Reisfeld B., Mayeno A. (eds) Computational Toxicology. *Methods in Molecular Biology (Methods and Protocols)*, Vol. 929. Humana Press, Totowa, NJ

Clewell, H.J. (2005). Use of mode of action in risk assessment: past, present, and future. *Reg Tox Pharm*, Vol. 42, pp. 3-14

Clewell HJ, Andersen ME. (1996). Use of physiologically-based pharmacokinetic modeling to investigate individual versus population risk. *Toxicology*, Vol. 111, pp. 315-329.

Clewell, H. J., III, and Andersen, M. E. (1985). Risk assessment extrapolations and physiological modeling. *Toxicol. Ind. Health*, Vol. 1, pp. 111-131

Clewell, R.A. and Clewell, H.J. (2008). Development and Specification of Physiologically Based Pharmacokinetic Models for Use in Risk Assessment. *Reg. Tox. Pharm.*, Vol. 50, pp. 129-143.

Clewell, H.J., Andersen, H.J., and Blaauboer, B.J. (2008). On the incorporation of chemical-specific information in risk assessment. *Toxicology Letters*, Vol. 180, pp. 100-109

Clewell, H.J., Gentry, P.R., Covington, T.R., Sarangapani, R., and Teeguarden, J.G. (2004). Evaluation of the potential impact of age- and gender-specific pharmacokinetic differences on tissue dosimetry. *Toxicol. Sci.*, Vol. 79, pp. 381-393.

Clewell, HJ, Andersen, ME, and Barton, HA. (2002). A consistent approach for the application of pharmacokinetic modeling in cancer and noncancer risk assessment. *Environ Health Perspect*, Vol. 110, pp. 85- 93.

Clewell HJ, Lee T, Carpenter, RL. (1994). Sensitivity of physiologically-based pharmacokinetic models to variation in model parameters: methylene chloride. *Risk Anal*, Vol. 14, pp. 521-531.

Chavan, S. *et al*. (2017). Combining *in Vivo*, *in Vitro* and toxicokinetics data in readacross: A case study using caffeine (SOT). Presented at SOT 2017, Baltimore, MD, March 12 - 16, 2017. https://doi.org/10.23645/epacomptox.5178604

Christian MS, Brent RL (2001). Teratogen update: evaluation of the reproductive and developmental risks of caffeine. *Teratology*, Vol. 64, No. 1, pp. 51-78.

CIR (2018). Safety Assessment of Methylxanthines as Used in Cosmetics. Tentative Report for Public Comment, release date: October 5, 2018. Cosmetic Ingredient Review, Washington, DC.

Cornelis MC, El-Sohemy A, Campos H. (2007). Genetic polymorphism of the adenosine A2A receptor is associated with habitual caffeine consumption. *Am J Clin Nutr*. Vol. 86, No.1, pp. 240-244

Cosmetics Europe (2017). Caffeine – skin penetration (frozen human skin) after 24h. TF ADME data, Sep 6, 2017

Daly JW (2007). Caffeine analogs: biomedical impact. *Cell Mol Life Sci*., Vol. 64. No. 16, pp. 2153-69.

Daly JW, Butts-Lamb P, Padgett W (1983). Subclasses of Adenosine Receptors in the Central Nervous System: Interaction with Caffeine and Related Methylxanthines. *Cellular and Molecular Neurobiology*, Vol. 3, No. 1, pp. 69-80

Denaro CP, Brown CR, Jacob P III, Benowitz NL (1991). Effectrs of caffeine with repeated dosing. *Eur J Clin Pharmacol*, Vol. 40, pp. 273-278.

Dias M, Farinha A, Faustino E, Hadgraft J, Pais J, Toscano C (1999). Topical delivery of caffeine from some commercial formulations. *International Journal of Pharmaceutics*, Vol. 182, pp. 41–47

Doucet O, Ferrero L, Garcia N, Zastrow L (1998). O/W emulsion and W/O/W multiple emulsion: physical characterization and Skin pharmacokinetic comparison in the delivery process of caffeine. *Int J Cosm Sci*, Vol. 20, pp. 283-295

Dudzina T, Delmaar CJE, Biesterbos JWH, Bakker MI, Bokkers BGH, Scheepers PTJ, van Engelen JGM, Hungerbuehler K, von Goetz N, (2015). The probabilistic aggregate consumer exposure model (PACEM): Validation and comparison to a lower-tier assessment for the cyclic siloxane D5. *Environment International*, Vol. 79, pp. 8–16

Drug Bank, consulted on March 27, 2019, caffeine query, https://www.drugbank.ca/drugs/DB00201

ECHA/EFSA (2018). Guidance for the identification of endocrine disruptors in the context of Regulations (EU) No 528/2012 and (EC) No 1107/2009. Prepublication version, 7 June 2018

EFSA (2015). Scientific Opinion on the safety of caffeine. *EFSA Journal*, Vol. 13, No. 5, pp. 4102, 120 pp. European Food Safety Authority, Parma, Italy. https://doi.org/10.2903/j.efsa.2015.4102.

Eilstein J, Grégoire S, Fabre A, Arbey E, Géniès C, Duplan H, Rothe H, Ellison C, Cubberley R, Schepky A, Lange D, Klaric M, Hewitt NJ, Jacques-Jamin C. (2020), Use of human liver and EpiSkin™ S9 subcellular fractions as a screening assay to compare the *in vitro* hepatic and dermal metabolism of 47 cosmetics-relevant chemicals. *JAT*, Vol. 40, No. 3, pp. 416-433. https://doi.org/10.1002/jat.3914.

Essayan DM (2001). Cyclic nucleotide phosphodiesterases. *J Allergy Clin Immunol*. No. 108, Vol. 5, pp. 671-80.

Faassen F, Vogel G, Spanings H, Vromans H (2003). Caco-2 permeability, P-glycoprotein transport ratios and brain penetration of heterocyclic drugs. *Int J Pharm,* Vol. 16, Vol. 263, No. 1-2, pp. 113-22.

Foukas LC, Daniele N, Ktori C, Anderson KE, Jensen J, Shepherd PR (2002). Direct effects of caffeine and theophylline on p110 delta and other phosphoinositide 3-kinases. Differential effects on lipid kinase and protein kinase activities. *J Biol Chem.,* Vol. 277, No. 40, pp. 37124-30

Gajewska M, Paini A, Sala Benito JV, Burton J, Worth A, Urani C, Briesen H, Schramm K-W (2015). *In vitro*-to-*in vivo* correlation of the skin penetration, liver clearance and hepatotoxicity of caffeine. *Food Chem Tox*, Vol. 75, pp. 39-49

Gajewska M, Worth A, Urani C, Briesen H, Schramm K-W (2014). Application of physiologically-based toxicokinetic modelling in oral-to-dermal extrapolation of threshold doses of cosmetic ingredients. *Toxicology Letters*, Vol. 227, pp. 189–202

Gracia-Lor E, Rousis N, Zuccato E, Bade R, Baz-Lomba J, Castrignanò E, Causanilles A, Hernández F, Kasprzyk-Hordern B, Kinyua J, McCall A, van Nuijs A, Plósz B, Ramin P, Ryu Y, Santos M, Thomas K, de Voogt P, Yang Z, Castiglioni S (2017). Estimation of caffeine intake from analysis of caffeine metabolites in wastewater. *Sci Total Environ*, Vol. 609, pp. 1582-1588

Géniès C, Jamin EL, Debrauwer L, Zalko D, Person EN, Eilstein J, Grégoire S, Schepky A, Lange D, Ellison C, Roe A, Salhi S, Cubberley R, Hewitt NJ, 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.* Vol. 39, pp. 385-397

Gerlowski, LE, and Jain, RJ. (1983). Physiologically based pharmacokinetic modeling: Principles and applications. *J Pharm. Sci.* Vol. 72, pp. 1103-1126

Ginsberg G, Hattis D, Russ A, Sonawane B (2004). Physiologically based pharmacokinetic (PBPK) modeling of caffeine and theophylline in neonates and adults: implications for assessing children's risks from environmental agents. *Journal of Toxicology and Environmental Health (Part A)*, Vol. 67, pp. 297– 329

Grosso G, Godos J, Galvano F, Giovannucci EL (2017). Coffee, Caffeine, and Health Outcomes: An Umbrella Review. *Annual Review of Nutrition*, Vol. 37, pp. 131-156

Hawke TJ, Allen DG, Lindinger MI (2000). Paraxanthine, a caffeine metabolite, dose dependently increases Ca^{2+} in skeletal muscle. *J Appl Physiol*, Vol. 89, pp. 2312–2317

Herman A, Herman AP (2013). Caffeine's Mechanisms of Action and Its Cosmetic Use. *Skin Pharmacol Physiol*, Vol. 26, No. 8–14

Hewitt NJ, Grégoire, Cubberley, Duplan, Eilstein, Ellison, Lester, Fabian, Fernandez, Géniès, Jacques-Jamin, Klaric, Rothe, Sorrell, Lange, Schepky (2019). Measurement of the penetration of 56 cosmetic relevant chemicals into and through human skin using a standardized protocol, *JAT*, Vol. 40, No. 30, pp. 403-415.

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

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

Kong, Jones, Koop, Zhang, Duff, and Chen (2008). Caffeine Induces Ca^{2+} Release by Reducing The Threshold for Luminal Ca^{2+} Activation of the Ryanodine Receptor. *Biochem J*., Vol. 414, No. 3, pp. 441–452

Kopecna M, Macháček M, Prchalová E, Štěpánek P, Drašar P, Kotora M, Vávrová K (2017). Dodecyl Amino Glucoside Enhances Transdermal and Topical Drug Delivery via reversible Interaction with Skin Barrier Lipids. *Pharm Res*, Vol. 34, pp. 640-653.

Kot M, Daniel WA (2008). Caffeine as a marker substrate for testing cytochrome P450 activity in human and rat. *Pharmacological Reports*, Vol. 60, pp. 789–797.

Lave, T., Dupin, S., Schmitt, C., Chou, R.C., Jaeck, D., Coassolo, P. (1997). Integration of *in vitro* data into allometric scaling to predict hepatic metabolic clearance in man: application to 10 extensively metabolized drugs. *J. Pharm. Sci*., Vol. 86, No. 5, pp. 584-590.

Lelo, A., Birkett, D.J., Robson, R.A., Miners, J.O. (1986). Comparative pharmacokinetics of caffeine and its primary demethylated metabolites paraxanthine, theobromine and theophylline in man. *Br. J. Clin. Pharmacol*. Vol. 22, pp. 177–182.

Lindstedt, SL and Schaeffer, PJ. (2001). Use of allometry in predicting anatomical and physiological parameters of mammals. *Laboratory Animals*, Vol. 36, pp. 1–19

Livingston EH and Lee S, (2001). Body surface area prediction in normal-weight and obese patients. *Am J Physiol Endocrinol Metab*, Vol. 281, pp. E586-E591.

Luo L, Lane ME (2015). Topical and transdermal delivery of caffeine. *International Journal of Pharmaceutics*, Vol. 490, pp. 155–164

McNamara DR, Golden D, Gibney M, Hall B, Tozer S, Safford B, Coroama M, Leneveu-Duchemin MC, Steiling W (2007). Probabilistic modelling of European consumer exposure to cosmetic products. *Food and Chemical Toxicology*, Vol. 45, pp. 2086–2096

Metro D, Cernaron V, Santoro D, Papa M, Buemi M, Benvenga S, Manasser L (2017). Beneficial effects of oral pure caffeine on oxidative stress. *Journal of Clinical & Translational Endocrinology*, Vol. 10, pp. 22–27

Monteiro JP, Alves MG, Oliveira PF, Silva BM (2016). Structure-Bioactivity Relationships of Trying to Make Sense of All the Promises and the Drawbacks. *Molecules*, Vol. 21, pp. 974

Muller CE, Jacobson KA (2011). Xanthines as adenosine receptor antagonists. *Handb. Exp. Pharmacol*., Vol. 200, pp. 151–199

Nehlig A, Daval J-L, Debry G. (1992). Caffeine and the central nervous system: mechanisms of action, biochemical, metabolic and psychostimulant effects. *Brain Research Reviews*, Vol. 17, pp. 139-170

NTP (1998). Tech. Rep. No. 473, Theophylline. National Toxicology Program, US. Department of Health and Human Services, *NIH-Publication* No. 98-3963; NTIS, PB99-113342, Research Triangle Park, NC.

Nyska A, Herbert RA, Chan PC, Haseman JK, Hailey JR (1998). Theophylline-induced mesenteric periarteritis in F344/N rats. *Arch Toxicol*, Vol. 72, No. 11, pp. 731–737

OECD (2017). Chemical Safety Assessment Workflow Based on Exposure Considerations and Non-Animal Methods. Series on Testing & Assessment No. 275, ENV/JM/MONO(2017)27, OECD,

Paris. Available at http://www.oecd.org/chemicalsafety/testing/series-testing-assessment-publications-number.htm.

OECD (2002). CAFFEINE, CAS: 58-08-2. SIDS Initial Assessment Report for SIAM 14, OECD.

OECD (2001). Theophylline, CAS 58-55-9. SIDS Initial Assessment Report for SIAM 13, OECD.

Otberg N, Patzelt A, Rasulev U, Hagemeister T, Linscheid M, Sinkgraven R, Sterry W, Lademann J (2008). The role of hair follicles in the percutaneous absorption of caffeine. *Br J Clin Pharmacol*, Vol. 65, pp. 488–492

Patlewicz G, Jeliazkova N, Safford RJ, Worth AP, Aleksiev B (2008). An evaluation of the implementation of the Cramer classification scheme in the Toxtree software. *SAR and QSAR in Environmental Research*, Vol. 19, No. 5-6, pp. 495-524

Peyret, T., Poulin, P., Krishnan, K. (2010). A unified algorithm for predicting partition coefficients for PBPK modeling of drugs and environmental chemicals. Toxicol. *Appl. Pharmacol,* Vol. 249, No.3, pp. 197–207

Preissner S, Kroll K, Dunkel M, Senger C, Goldsobel G, Kuzman D, Guenther S, Winnenburg R, Schroeder M, Preissner R (2010). SuperCYP: a comprehensive database on Cytochrome P450 enzymes including a tool for analysis of CYP-drug interactions. *Nucleic Acids Res.*, Vol. 38(Database issue), D237-43

PubChem, consulted on March 27, 2019, caffeine query by CASRN 58-08-2, https://pubchem.ncbi.nlm.nih.gov/compound/caffeine

Sarkaria JN, Busby EC, Tibbetts RS, Roos P, Taya Y, Karnitz LM, Abraham RT (1999). Inhibition of ATM and ATR kinase activities by the radiosensitizing agent, caffeine. *Cancer Res.*, Vol. 59, No. 17, pp. 4375-82

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

Schmitt, W. (2008). General approach for the calculation of tissue to plasma partition coefficients. *Toxicol. Vitr.* Vol. 22, pp. 457–467.

Schultz TW, Richarz A-N, Cronin MTD (2019). Assessing uncertainty in read-across: Questions to evaluate toxicity predictions based on knowledge gained from case studies. *Computational Toxicology*, Vol. 9, No. 1–11

Schultz TW, Amcoff P, Berggren E, Gautier F, Klaric M, Knight DJ, Mahony C, Schwarz M, White A, Cronin MTD (2015). A strategy for structuring and reporting a read-across prediction of toxicity. *Regulatory Toxicology and Pharmacology*, Vol. 72, pp. 586–601

Sengpiel V, Elind E, Bacelis J, Nilsson S, Grove J, Myhre R, Haugen M, Meltzer HM, Alexander J, Jacobsson B, Brantsaeter AL (2013). Maternal caffeine intake during pregnancy is associated with birth weight but not with gestational length: results from a large prospective observational cohort study. *BMC Med.,* Vol. 19, pp. 11-42.

Shibata M, Wachi M, Kawaguchi M, Kojima J, Onodera K (2000). Teratogenic and fetal toxicity following intravenous theophylline administration in pregnant rabbits is related to maternal drug plasma levels. *Methods Find Exp Clin Pharmacol*, Vol. 22, pp. 101–107

Sipes NS, Martin MT, Kothiya P, Reif DM, Judson RS, Richard AM, Houck KA, Dix DJ, Kavlock RJ, and Knudsen TB (2013). Profiling 976 ToxCast Chemicals across 331 Enzymatic and Receptor Signaling Assays. *Chem. Res. Toxicol.*, Vol. 26, pp. 878−895

Tarka SM Jr (1982). The toxicology of cocoa and methylxanthines: a review of the literature. *Crit Rev Toxicol,* Vol. 9, No. 4, pp. 275-312; cited in: Theocorp Holding Company (2010). GRAS Exemption claim for Theobromine

Theocorp Holding Company (2010). GRAS Exemption claim for Theobromine, Date of filing: May 17, 2010, GRN No. 340. Theocorp Holding Company, LLC, Metairie, LA.

Toner F, Garcia B, Roper C, Madden S (2009). Effect of Repeat Application of [14C]-Caffeine on Percutaneous Absorption Through Human Skin *in vitro*. Poster, Occupational and Environmental Exposures of Skin to Chemicals – 2009 (OEESC), June 14-17, 2009, Edinburgh, UK

Trauer S, Patzelt A, Otberg N, Knorr F, Rozycki C, Balizs G, Büttemeyer R, Linscheid M, Liebsch M, Lademann J (2009). Permeation of topically applied caffeine through human skin – a comparison of *in vivo* and *in vitro* data. *Br J Clin Pharmacol*, Vol. 68, No. 2, pp. 181–186

Troutman JA, Rick DL, Stuard SB, Fisher J, Bartels MJ (2015). Development of a physiologically-based pharmacokinetic model of 2-phenoxyethanol and its metabolite phenoxyacetic acid in rats and humans to address toxicokinetic uncertainty in risk assessment. *Reg Tox Pharm*, Vol. 73, pp. 530-543.

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

U.S. EPA (2006). Approaches for the Application of Physiologically Based Pharmacokinetic (PBPK) Models and Supporting Data in Risk Assessment. National Center for Environmental Assessment, United States Environmental Protection Agency, Washington, DC; EPA/600/R-05/043F. Available from: National Technical Information Service, Springfield, VA, and online at http://epa.gov/ncea

U.S. EPA (1994). Methods for Derivation of Inhalation Reference Concentrations and Application of Inhalation Dosimetry, EPA/600/8-90/066F. United States Environmental Protection Agency, Research Triangle Park, NC.

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

Wang S.J. (2007). Caffeine facilitation of glutamate release from rat cerebral cortex nerve terminals (synaptosomes) through activation protein kinase C pathway: an interaction with presynaptic adenosine A1 receptors. *Synapse*, Vol. 61, No. 6, pp. 401-11

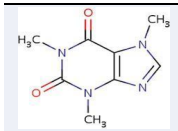
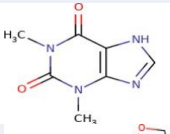
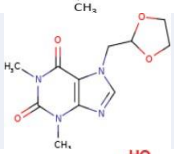
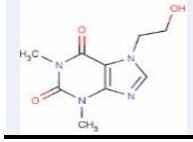
WHO (2010). Characterization and Application of Physiologically Based Phamacokinetic Models in Risk Assessment. World Health Organization, Geneva. https://apps.who.int/iris/handle/10665/44495.

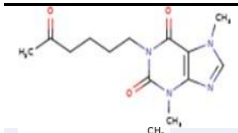
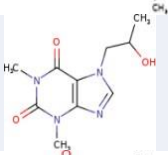
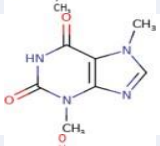
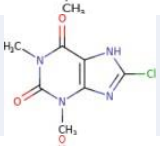
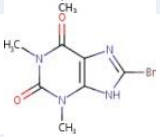
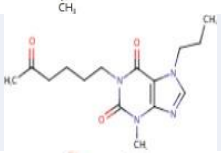
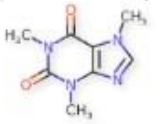
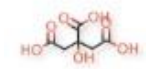
Yoon, M., J. L. Campbell, M. E. Andersen, and H. J. Clewell (2012), Quantitative in vitro to in vivo Extrapolation of Cell-Based Toxicity Assay Results. Crit Rev Toxicol, Vol. 42, No. 8, pp. 633-52. https://doi.org/10.3109/10408444.2012.692115.

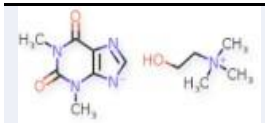
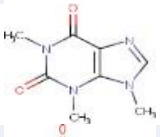
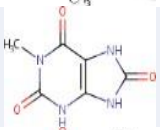
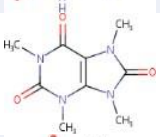
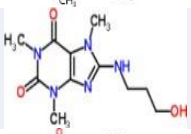
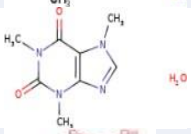
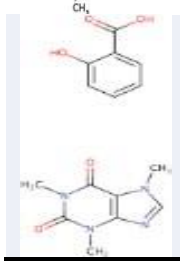
York RG, Randall JL, Scott Jr WJ (1986). Teratogenicity of paraxanthine (1,7-dimethylxanthine) in C57BL/6J mice. *Teratology*, Vol. 34, pp. 279-282

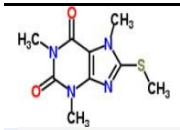
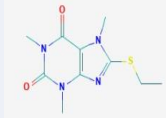
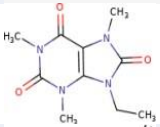
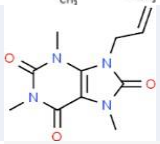
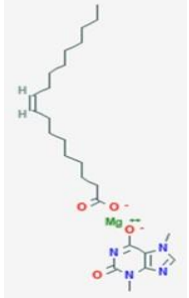
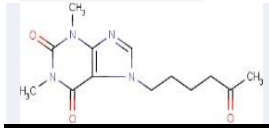
Zesch A, Schaefer H, Stüttgen G (1979). The quantitative distribution of percutaneously applied caffeine in the human skin. *Arch Dermatol Res*. Vol. 266, No. 3, pp. 277-283

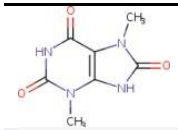
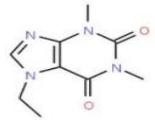
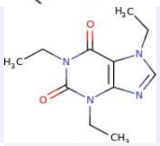
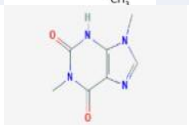
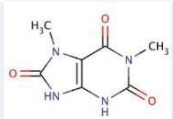
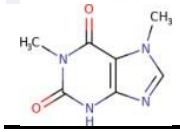
Annex 1. Analogues of caffeine based on CE-TOXGPS

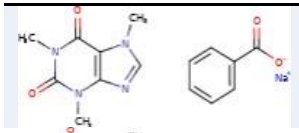
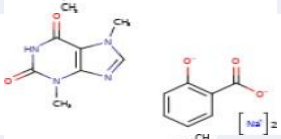
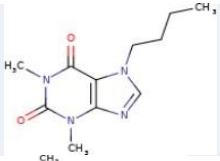
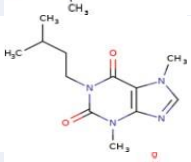
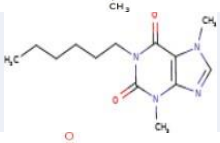
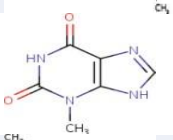
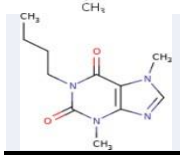
Compound Summary Structure & IDs CMS ID	Data Summary	Data Summary Registry Numbers List	Data Summary Studies Count	Fingerprints MACCS Fingerprint Tanimoto	Fingerprints RDKit MolFingerprint Tanimoto
	CAFFEINE 227 studies	58-08-2 (Active)	227	1	1
	THEOPHYLLINE One study	58-55-9 (Active)	1	0,949152542	0,918719212
	DOXOFYLLINE	69975-86-6 (Active)	0	0,835820896	0,838842975
	DYPHYLLINE	479-18-5 (Active)	0	0,788732394	0,844074844
	ETOFYLLINE	519-37-9 (Active)	0	0,835820896	0,902222222

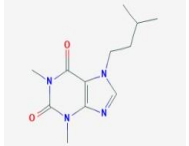
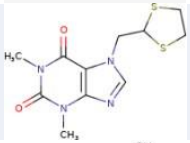
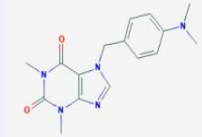
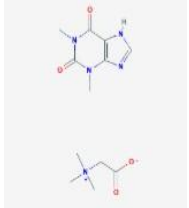
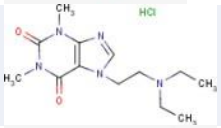
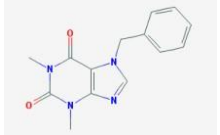
	PENTOXIFYLLINE	6493-05-6 (Active)	0	0,863636364	0,825203252
	PROXYPHYLLINE	603-00-9 (Active)	0	0,876923077	0,861995754
	THEOBROMINE 46 studies	83-67-0 (Active)	46	0,9	0,92364532
	8-CHLOROTHEOPHYLLINE	85-18-7 (Active)	0	0,888888889	0,833333333
	8-BROMOTHEOPHYLLINE	10381-75-6 (Active)	0	0,888888889	0,820960699
	PROPENTOPHYLLINE	55242-55-2 (Active)	0	0,850746269	0,810379242
	CAFFEINE CITRATE	69-22-7 (Active)	0	0,77027027	0,833675565
					

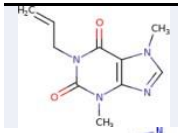
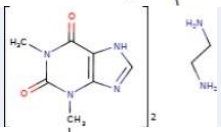
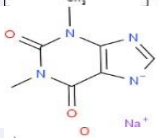
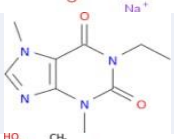
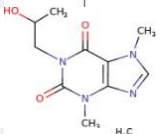
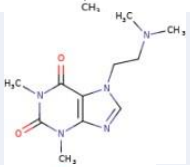
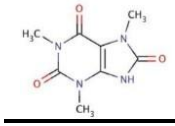
	OXTRIPHYLLINE	4499-40-5 (Active)	0	0,730769231	0,886524823
	ISOCAFFEINE	519-32-4 (Active)	0	1	0,90625
	3-METHYLURIC ACID	605-99-2 (Active)	0	0,686567164	0,816666667
	1,3,7,9-TETRAMETHYLURIC ACID	2309-49-1 (Active)	0	0,85	0,950819672
	1H-PURINE-2,6-DIONE, 3,7- DIHYDRO-1,3,7-TRIMETHYL-8-(METHYLAMINO)-CAFFEINE	5422-30-0 (Active)	0	0,919354839	0,816901408
	CAFFEINE MONOHYDRATE	5743-12-4 (Active)	0	0,93442623	1
	CAFFEINE SALICYLATE	5743-22-6 (Active)	0	0,802816901	0,813627255

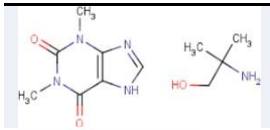
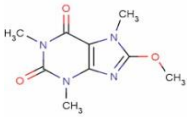
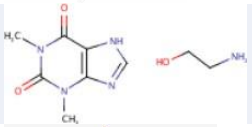
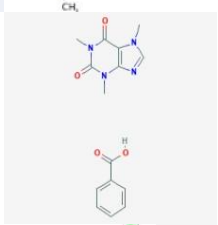
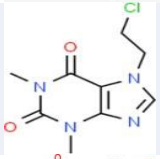
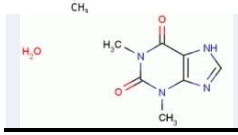
	8-(METHYLTHIO)-CAFFEINE	6287-54-3 (Active)	0	0,919354839	0,815261044
	8-(ETHYLTHIO)-CAFFEINE	6287-57-6 (Active)	0	0,890625	0,803960396
	URIC ACID, 9-ETHYL-1,3,7-TRIMETHYL-	6287-59-8 (Active)	0	0,80952381	0,863829787
	URIC ACID, 9-ALLYL-1,3,7- TRIMETHYL	6287-61-2 (Active)	0	0,836065574	0,838842975
	THEOBROMINE MAGNESIUM OLEATE	6767-73-3 (Active)	0	0,791666667	0,84
	PURIN-2,6-DIONE, 1,2,3,6-TETRAHYDRO-1,3-DIMETHYL- 7-(5-OXOHEXYL)-	10226-54-7 (Active)	0	0,863636364	0,854736842

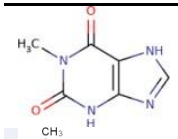
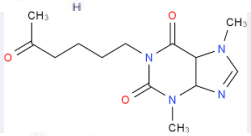
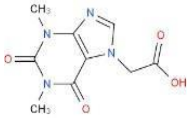
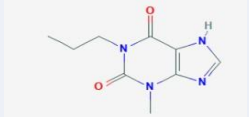
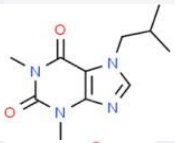
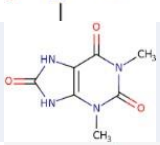
	3,7-DIMETHYLURIC ACID	13087-49-5 (Active)	0	0,738461538	0,897619048
	TRICLOFYLLINE [INN:DCF]	17243-70-8 (Active)	0	0,777777778	0,833675565
	7-ETHYLTHEOPHYLLINE	23043-88-1 (Active)	0	0,95	0,933333333
	XANTHINE, 1,3,7-TRIETHYL-	31542-50-4 (Active)	0	0,901639344	0,876889849
	1,9-DIMETHYLXANTHINE	33073-01-7 (Active)	0	0,915254237	0,878345499
	1,7-DIMETHYLURIC ACID	33868-03-0 (Active)	0	0,738461538	0,873205742
	1,7-DIMETHYLXANTHINE	611-59-6 (Active)	0	0,9	0,899014778

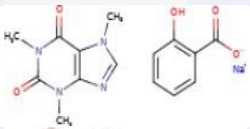
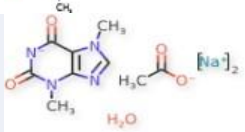
	CAFFEINE SODIUM BENZOATE	8000-95-1 (Active)	0	0,850746269	0,88453159
	THEOBROMINE SODIUM SALICYLATE	8048-31-5 (Active)	0	0,857142857	0,92364532
	7-BUTYLTHEOPHYLLINE	1021-65-4 (Active)	0	0,890625	0,888402626
	1-ISOPENTYLTHEOBROMINE	1024-65-3 (Active)	0	0,904761905	0,831967213
	1-HEXYL-3,7- DIMETHYLXANTHINE	1028-33-7 (Active)	0	0,890625	0,849372385
	3-METHYLXANTHINE	1076-22-8 (Active)	0	0,838709677	0,834975369
	1-BUTYLTHEOBROMINE	1143-30-2 (Active)	0	0,890625	0,854736842

	7-ISOAMYLTHEOPHYLLINE	1146-79-8 (Active)	0	0,890625	0,882608696
	NESTIFYLLINE	116763-36-1 (Active)	0	0,903225806	0,835390947
	3,7-DIMETHYL-1-PROPARGYLXANTHINE	14114-46-6 (Active)	0	0,965517241	0,87311828
	DIMABEFYLLINE	1703-48-6 (Active)	0	0,890625	0,805555556
	THEOPHYLLINE-BETAINE	17140-27-1 (Active)	0	0,814285714	0,858447489
	7-DIETHYLAMINOETHYL THEOOHYLLINE HYDROCHLORIDE	17140-68-0 (Active)	0	0,838235294	0,87311828
	7-BENZYLTHEOPHYLLINE	1807-85-8 (Active)	0	0,933333333	0,860169492

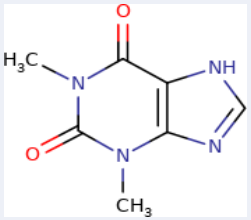
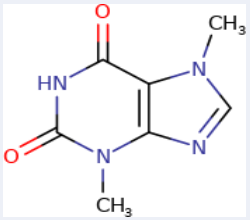
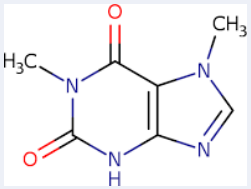
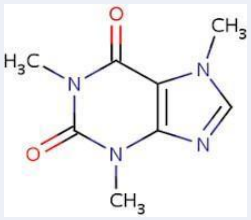
	1-ALLYLTHEOBROMINE	2530-99-6 (Active)	0	0,948275862	0,871244635
	ETAMIPHYLLIN	314-35-2 (Active)	0	0,890625	0,87311828
	AMINOPHYLLINE	317-34-0 (Active)	0	0,802816901	0,905339806
	1,3-DIMETHYLXANTHINE SODIUM SALT (THEOPHYLLINE SODIUM SALT)	3485-82-3 (Active)	0	0,933333333	0,918719212
	1-ETHYLTHEOBROMINE	39832-36-5 (Active)	0	0,95	0,908277405
	PROTHEOBROMINE	50-39-5 (Active)	0	0,876923077	0,821862348
	DIMETHAZAN	519-30-2 (Active)	0	0,890625	0,886462882
	1,3,7-TRIMETHYLURIC ACID	5415-44-1 (Active)	0	0,80952381	0,966666667

	AMBUPHYLLINE	5634-34-4 (Active)	0	0,76	0,879716981
	8-METHOXYCAFFEINE	569-34-6 (Active)	0	0,919354839	0,813627255
	THEOPHYLLINE OLAMINE	573-41-1 (Active)	0	0,767123288	0,898795181
	CAFFEINE BENZOATE	5743-17-9 (Active)	0	0,850746269	0,88453159
	7- (CHLOROETHYL)THEOPHYLLINE	5878-61-5 (Active)	0	0,903225806	0,90625
	CAFFEINE HYDROCHLORIDE	5892-18-2 (Active)	0	0,93442623	1
	THEOPHYLLINE MONOHYDRATE	5967-84-0 (Active)	0	0,875	0,918719212

	1-METHYLXANTHINE	6136-37-4 (Active)	0	0,825396825	0,810344828
	LISOPHYLLINE	6493-06-7 (Active)	0	0,791666667	0,831967213
	7-THEOPHYLLINEACETIC ACID	652-37-9 (Active)	0	0,861538462	0,867521368
	3-METHYL-1- PROPYLXANTHINE	72117-78-3 (Active)	0	0,875	0,810344828
	ISBUFYLLINE	90162-60-0 (Active)	0	0,95	0,894273128
	1,3-DIMETHYLURIC ACID	944-73-0 (Active)	0	0,757575758	0,895238095
	AMINOPHYLLINE HYDRATE	72487-55-9 (Active)	0	0,743243243	0,910194175

	THEOPHYLLINE SODIUM GLYCINATE [USAN]	8000-10-0 (Active)	0	0,8	0,886255924
	CAFFEINE SODIUM SALICYLATE	8002-85-5 (Active)	0	0,780821918	0,813627255
	THEOBROMINE SODIUM ACETATE	8002-88-8 (Active)	0	0,763888889	0,906024096

Annex 2. QSAR Toolbox v 4.3 profilers for caffeine, theobromine, theophylline and paraxanthine

Substance identity	Analogue 1	Analogue 2	Analogue 3	Target substance
Structure				
CAS number	58-55-9	83-67-0	611-59-6	58-08-2
Chemical name	Theophyline	Theobromine	1,7-Dimethylxanthine	Caffeine
Other identifier			Paraxanthine	
SMILES	CN1C(=O)N(C)c2nc[nH]c2C1=O	CN1C(=O)NC(=O)c2c1ncn2C	CN1C(=O)Nc2ncn(C)c2C1=O	CN1C(=O)N(C)c2ncn(C)c2C1=O
Profilers				
General Mechanistic				
DNA binding by OASIS	No alert found	No alert found	No alert found	No alert found
DNA binding by OECD	SN1 >> Iminium Ion Formation >> Aliphatic tertiary amines	No alert found	SN1 >> Iminium Ion Formation >> Aliphatic tertiary amines	SN1 >> Iminium Ion Formation >> Aliphatic tertiary amines
Toxic hazard classification by Cramer	High (Class III)	High (Class III)	High (Class III)	High (Class III)
Protein binding by OECD	Acylation >> Direct Acylation Involving a Leaving group >> Acetates	Acylation >> Direct Acylation Involving a Leaving group >> Acetates	Acylation >> Direct Acylation Involving a Leaving group >> Acetates	Acylation >> Direct Acylation Involving a Leaving group >> Acetates
Estrogen Receptor Binding	Non binder, without OH or NH2 group	Non binder, without OH or NH2 group	Non binder, without OH or NH2 group	Non binder, without OH or NH2 group
Protein binding by OASIS	No alert found	No alert found	No alert found	No alert found
Toxic hazard classification by Cramer (extended)	High (Class III)	High (Class III)	High (Class III)	High (Class III)
Endpoint Specific				
Skin irritation/corrosion Exclusion rules by BfR	Group All Melting Point > 200 C; Group CN Melting Point > 180 C;	Group All Melting Point >200 C; Group CN Melting Point >180 C;	Group All Melting Point > 200 C; Group CN Melting Point > 180 C;	Group All Melting Point >200 C; Group CN Melting Point >180 C;

	Group CN Vapour Pressure < 0.001 Pa; Undefined	Group CN Vapour Pressure < 0.001 Pa; Undefined	Group CN Vapour Pressure < 0.001 Pa; Undefined	Group CN Vapour Pressure < 0.001 Pa; Undefined
Oncologic Primary Classification	Not classified	Not classified	Not classified	Not classified
Acute aquatic toxicity classification by Verhaar (Modified)	Class 5 (Not possible to classify according to these rules)	Class 5 (Not possible to classify according to these rules)	Class 5 (Not possible to classify according to these rules)	Class 5 (Not possible to classify according to these rules)
Eye irritation/corrosion Exclusion rules by BfR	Group All Melting Point > 200 C; Undefined	Group All Melting Point > 200 C; Undefined	Group All Melting Point > 200 C; Undefined	Group All Melting Point > 200 C; Undefined
DNA alerts for AMES by OASIS	No alert found	No alert found	No alert found	No alert found
Acute aquatic toxicity MOA by OASIS	Reactive unspecified	Reactive unspecified	Reactive unspecified	Reactive unspecified
rtER Expert System - USEPA	No alert found	No alert found	No alert found	No alert found
DNA alerts for CA and MNT by OASIS	No alert found	No alert found	No alert found	No alert found
Keratinocyte gene expression	High gene expression >> N- Acylamides	High gene expression >> N- Acylamides	High gene expression >> N- Acylamides	High gene expression >> N- Acylamides
DART scheme	Known precedent reproductive and developmental toxic potential; Purine and pyrimidine-like derivatives (7b)	Known precedent reproductive and developmental toxic potential; Purine and pyrimidine-like derivatives (7b)	Known precedent reproductive and developmental toxic potential; Purine and pyrimidine-like derivatives (7b)	Known precedent reproductive and developmental toxic potential; Purine and pyrimidine-like derivatives (7b)
Skin irritation/corrosion Inclusion rules by BfR	Inclusion rules not met	Inclusion rules not met	Inclusion rules not met	Inclusion rules not met
Aquatic toxicity classification by ECOSAR	Carbonyl Ureas; Imidazoles	Carbonyl Ureas; Imidazoles	Carbonyl Ureas; Imidazoles	Carbonyl Ureas; Imidazoles
in vitro mutagenicity (Ames test) alerts by ISS	No alert found	No alert found	No alert found	No alert found
Carcinogenicity (genotox and nongenotox) alerts by ISS	Structural alert for nongenotoxic carcinogenicity; Imidazole, benzimidazole (Nongenotox)	No alert found	No alert found	No alert found
Respiratory sensitisation	No alert found	No alert found	No alert found	No alert found
Retinoic Acid Receptor Binding	Not possible to classify according to these rules	Not possible to classify according to these rules	Not possible to classify according to these rules	Not possible to classify according to these rules
Protein binding alerts for Chromosomal aberration by OASIS	AN2 >> Michael type addition to activated double bond of pyrimidine bases >> Pyrimidines and Purines; AN2 >> Shiff base formation with carbonyl group of pyrimidine or purine bases >> Pyrimidines and Purines	No alert found	No alert found	AN2 >> Michael type addition to activated double bond of pyrimidine bases >> Pyrimidines and Purines; AN2 >> Shiff base formation with carbonyl group of pyrimidine or purine bases >> Pyrimidines and Purines

in vivo mutagenicity (Micronucleus) alerts by ISS	H-acceptor-path3-H-acceptor	H-acceptor-path3-H-acceptor	H-acceptor-path3-H-acceptor	H-acceptor-path3-H-acceptor
Eye irritation/corrosion Inclusion rules by BfR	Inclusion rules not met	Inclusion rules not met	Inclusion rules not met	Inclusion rules not met

Annex 3. Summary data for repeat-dose and reproductive toxicity studies on theophylline, theobromine and paraxanthine

Theophylline

Test substance	Theophylline	
Reference	NTP (1998), OECD (2001)	
Study type	Repeated dose toxicity study (14-weeks)	
Species	Rat (F344/N)	
Route	Oral (diet)	
No. per group	10/sex	
Dose level	0, 1000, 2000, 4000 ppm (male: ca. 75, 125, 250 mg/kg bw/d; female: ca. 75, 125, 275 mg/kg bw/d)	
OECD GL	- (NTP design)	
Klimisch score	1 (Reliable without restriction)	
	Male	*Female*
NOAEL	LOAEL 1000 ppm (75 mg/kg bw/d)	LOAEL 1000 ppm (75 mg/kg bw/d)
Mortality	0/10	0/10
Body weight	No adverse impact	No adverse impact, ↑ mean body weight gain at 150 mg/kg
Haematology	2000, 4000 ppm: ↑ Mean cell volume, ↑ mean cell haemoglobin 4000 ppm: ↑ platelet count	1000, 2000, 4000 ppm: ↑ segmented neutrophil counts
Organ weights	4000 ppm: ↑ abs + rel kidney weights	4000 ppm: ↑ abs + rel lung weights
Histopathological findings	0, 1000 ppm: Minimal nephropathy (10/10, 10/10) 2000 ppm: mild nephropathy (10/10), mesenteric and pancreatic periarteritis (2/10) 4000 ppm: moderate nephropathy (10/10), mesenteric and pancreatic periarteritis (3/10)	1000, 2000 ppm: mesenteric and pancreatic periarteritis (1/10) 4000 ppm: mesenteric and pancreatic periarteritis (5/10)

Test substance	Theophylline	
Reference	NTP (1998), OECD (2001)	
Study type	Repeated dose toxicity study (14-weeks)	
Species	Mouse (B6C3F1)	
Route	Oral (diet)	
No. per group	10/sex	
Dose level	0, 1000, 2000, 4000 ppm (male: ca. 175, 400, 800 mg/kg bw/d; female: ca. 225, 425, 850 mg/kg/d bw)	
OECD GL	- (NTP design)	
Klimisch score	1 (Reliable without restriction)	
	Male	*Female*
NOAEL	LOAEL 1000 ppm (75 mg/kg bw/d) (based on reduced bw)	LOAEL 1000 ppm (225 mg/kg bw/d) (based on reduced bw)
Mortality	0/10	0/10
Body weight	1000, 2000, 4000 ppm: ↓ body weight and body weight gain	1000, 2000, 4000 ppm: ↓ body weight and body weight gain
Haematology	4000 ppm: ↑ Leukocyte, segmented neutrophil, lymphocyte count	2000, 4000 ppm: ↑ Leukocyte, segmented neutrophil counts
Organ weights	No relevant findings	2000, 4000 ppm: ↓ absolute and relative thymus weights
Histopathological findings	No relevant findings	No relevant findings

Test substance	Theophylline	
Reference	NTP (1998), OECD (2001)	
Study type	Repeated dose toxicity study (14-weeks)	
Species	Rat (F344/N)	
Route	Oral (gavage)	
No. per group	10/sex	
Vehicle	Corn oil	
Dose level	0, 37.5, 75, 150 mg/kg bw/d	
OECD GL	- (NTP design)	
Klimisch score	1 (Reliable without restriction)	
	Male	*Female*
NOAEL	Not determined by NTP, LOAEL 37.5 mg/kg bw/d	Not determined by NTP, LOAEL 37.5 mg/kg bw/d
Mortality	1/10 (150 mg/kg bw/d)	1/10 (150 mg/kg bw/d)
Body weight	No adverse impact	No adverse impact, ↑ mean body weight gain at 150 mg/kg
Haematology	37.5, 75, 150 mg/kg: ↑ mean cell haemoglobin 150 mg/kg bw/d: ↑ Mean cell volume	No relevant findings
Organ weights	150 mg/kg bw/d: ↓ thymus weights	150 mg/kg bw/d: ↓ thymus weights, ↑ abs + rel liver weights
Histopathological findings	slight dose-dependent increase in the incidence of periarteritis of the small- to medium-sized arteries adjacent to the mesenteric lymph nodes (1/10, 1/10, 2/10, 5/10)	slight dose-dependent increase in the incidence of periarteritis of the small- to medium-sized arteries adjacent to the mesenteric lymph nodes (0/10, 2/10, 2/10, 3/10)

Test substance	Theophylline	
Reference	NTP (1998), OECD (2001)	
Study type	Repeated dose toxicity study (14-weeks)	
Species	Mouse (B6C3F1)	
Route	Oral (gavage)	
No. per group	10/sex	
Vehicle	Corn oil	
Dose level	0, 75, 150, 300 mg/kg bw/d	
OECD GL	- (NTP design)	
Klimisch score	1 (Reliable without restriction)	
	Male	*Female*
NOAEL	NOAEL: 75 mg/kg bw/d (mortality, reduced bw)	NOAEL: 150 mg/kg bw/d (f) (mortality)
Mortality	1/10 (75 mg/kg bw/d), 3/10 (300 mg/kg bw/d)	10/10 (300 mg/kg bw/d)
Body weight	150, 300 mg/kg bw/d: reduced bw and bw gain	No relevant findings
Haematology	300 mg/kg: ↑ mean cell volume, mean cell haemoglobin	No relevant findings
Organ weights	No relevant findings	No relevant findings
Histopathological findings	No relevant findings	No relevant findings

Test substance	Theophylline	
Reference	NTP (1998), OECD (2001)	
Study type	Carcinogenicity study (2 years)	
Species	Rat (F344/N)	
Route	Oral (gavage)	
No. per group	50/sex	
Vehicle	Corn oil	
Dose level	0, 7.5, 25, 75 mg/kg bw/d	
OECD GL	NTP design (similar to OECD 451)	
Klimisch score	1 (Reliable without restriction)	
	Male	*Female*
Survival rate	23/50, 33/50, 29/50, 24/50	32/50, 30/50, 33/50, 33/50
Body weight	dose-related decrease in mean body weights	dose-related decrease in mean body weights
Histopathological findings		
Non-neoplastic effects	Mesenteric artery: chronic inflammation (2/50, 2/50, 3/50, 15/50)	None
Neoplastic effects	None	None
Decreased incidences	None	Mammary gland: fibroadenoma (22/50, 19/50, 12/50, 12/50); fibroadenoma or carcinoma (23/50, 20/50, 12/50, 12/50)
Level of evidence of Carcinogenic activity	No evidence	No evidence

Test substance	Theophylline
Reference	OECD (2001)
Study type	Reproductive toxicity - Developmental Toxicity / Teratogenicity
Species	Rat (Sprague-Dawley CD)
Route	Oral (diet)
Dose level	0, 1500, 3000, 4000 ppm = 0, 124, 218, 259 mg/kg bw/d
Treatment period	Gestation days (GD) 6 to 15
OECD GL	-
Klimisch score	2 (Reliable with restriction)
NOAEL	NOAEL maternal toxicity 1500 ppm (124 mg/kg bw/d) NOAEL developmental toxicity 1500 ppm (124 mg/kg bw/d)
Dams	4000 ppm: significant decrease in bw, bw gain and food consumption, clinical signs like piloerection and rough coat 3000 ppm: significant decrease in food consumption
Foetuses	4000 ppm: significant decrease in live foetuses per litter, significant decrease in average foetal body weight per litter 3000 ppm: significant decrease in average foetal body weight per litter No impact on the incidence of malformations (malformed foetuses per litter occurred with an incidence of 1.38%, 0.92%, 0.33%, and 1.57% for the vehicle control, low, medium, and high dose groups, respectively)

Test substance	Theophylline
Reference	OECD (2001)
Study type	Reproductive toxicity - Developmental Toxicity / Teratogenicity
Species	Mouse (Swiss CD-1)
Route	Oral (drinking water)
Dose level	0, 750, 1500, 2000 ppm = 0, 282, 372, 396 mg/kg bw/d
Treatment period	Gestation days (GD) 6 to 15
OECD GL	-
Klimisch score	2 (Reliable with restriction)
NOAEL	NOAEL maternal toxicity 750 ppm (282 mg/kg bw/d) NOAEL developmental toxicity 750 ppm (282 mg/kg bw/d)
Dams	2000 ppm: significant decrease in bw, bw gain and water consumption 1500 ppm: significant decrease in bw, bw gain and water consumption
Foetuses	2000 ppm: significant increase in in percentage resorptions per litter, decrease in average male and female foetal weight per litter 1500 ppm: significant increase in percentage resorptions per litter, decrease in average male and female foetal weight per litter 1500 and 2000 ppm: An increasing but statistically not significant trend for percentage malformed foetuses per litter

Test substance	Theophylline
Reference	OECD (2001)
Study type	Reproductive toxicity - Developmental Toxicity / Teratogenicity
Species	Rabbit (Kbl:JW), 20 females/group
Route	Intravenous (IV, using an automatic infusion pump)
Dose level	0, 15, 30 and 60 mg/kg bw/d (equivalent to maternal Cmax of 0, 30, 56 and 106 µg/mL, GD 6 and 18)
Treatment period	Gestation days (GD) 6 to 18
OECD GL	-
Klimisch score	2 (Reliable with restriction)
NOAEL	NOAEL maternal toxicity: 30 mg/kg bw/d NOAEL fetotoxicity/teratogenicity: 30 mg/kg bw/d
Dams	60 mg/kg bw/d: significant decrease in bw and food intake, clinical signs like accelerated respiration, abortion, sluggish startle reactions, dilatation of the auricular vessels and polyuria
Foetuses	60 mg/kg bw/d: Developmental toxicity such as an increased number of late deaths, increased foetal body weights (about 10% below concurrent controls) and effects on foetal morphology; increased rate of foetuses with cleft palates (8/103 foetuses in 2/14 litters) and with a 13th rib (63/103 foetuses)

Theobromine

Test substance	Theobromine	
Reference	Tarka (1982)	
Study type	Repeated dose toxicity study (13-weeks)	
Species	Rat (Sprague-Dawley)	
Route	Oral (diet)	
No. per group	10/sex	
Dose level	0, 200, 1000, 2000 ppm (0, 25, 125, 250 mg/kg bw/d)	
OECD GL	-	
Klimisch score	2 (Reliable with restriction)	
	Male	*Female*
NOAEL	Not determined, NOAEL defined at 125 mg/kg bw/d according to the results	Not determined, NOAEL defined at 250 mg/kg bw/d according to the results
Mortality	0/10	0/10
Body weight	2000 ppm: statistically significant reduction in bw gain and abs testicular weight	None
Haematology	None	None
Organ weights	4000 ppm: ↑ abs + rel kidney weights	None
Histopathological findings	No gross and/or histopathological lesions	No gross and/or histopathological lesions

Test substance	Theobromine
Reference	Theocorp Holding Company (2010)
Study type	Reproductive toxicity
Species	Rat (Sprague-Dawley)
Route	Oral (diet)
Dose level	0, 250 and 500 mg/kg bw/day
Treatment period	2 weeks and 4 weeks
OECD GL	-
Klimisch score	2 (Reliable with restriction)
NOAEL	< 250 mg/kg bw/d
Body weight	500 mg/kg bw/d: reduced bw gain
Organ weights	500 mg/kg bw/d: reduced weights of testicular and thymus tissue, relative prostate- and seminal vesicle weight
Histopathological findings	500 mg/kg bw/d: testicular toxicity after 2 and 4 weeks of dosing (degeneration and necrosis) 250 mg/kg bw/d: testicular toxicity after 4 weeks of dosing

Test substance	Theobromine
Reference	Theocorp Holding Company (2010)
Study type	Reproductive toxicity - Developmental Toxicity / Teratogenicity
Species	Rat (Sprague-Dawley)
Route	Oral (diet)
Dose level	0, 0.0625 or 0.135 % (equivalent to 53 or 99 mg/kg bw/d, at 99 mg/kg bw serum concentration 15-20 µg/mL)
Purity	>99.5 % purity (confirmed by HPLC/MS, gas-liquid chromatography, and infrared spectroscopy)
Treatment period	GD 6-19
OECD GL	-
Klimisch score	2 (Reliable with restriction)
NOAEL	NOAEL maternal toxicity: 53 mg/kg bw/d NOAEL fetotoxicity/teratogenicity: 99 mg/kg bw/d
Dams	99 mg/kg bw: significant reduction in food intake
Foetuses	No malformations occurred 99 mg/kg bw/d: slight decrease in fetal body weights, significant increase in a delay of osteogenesis (incompletely ossified or absent sternebrae and pubic bones). Effects are considered to be related to the significant reductions in food intake throughout GD 6-19.

Test substance	Theobromine
Reference	Theocorp Holding Company (2010)
Study type	Reproductive toxicity - Developmental Toxicity / Teratogenicity (GLP)
Species	Rabbit (New Zealand)
Route	Oral (gavage)
Dose level	0, 25, 75, 125, 200 mg/kg bw/d
Purity	>99.5 % purity (confirmed by HPLC/MS, gas-liquid chromatography, and infrared spectroscopy)
Treatment period	GD 6-29
OECD GL	-
Klimisch score	2 (Reliable with restriction)
NOAEL	NOAEL maternal toxicity: 75 mg/kg bw/d (24-86 µL/mL serum concentration) NOAEL fetotoxicity/teratogenicity: 25 mg/kg bw/day (based on incomplete ossification)
Dams	200 mg/kg bw/d: Mortality 40% 125 mg/kg bw/d: significant reduction in food intake
Foetuses	125, 200 mg/kg bw/d: decreases in foetal body weight and increases in various developmental variations 75, 125, 200 mg/kg bw/d: increased incidence of skeletal variations

Test substance	Theobromine
Reference	Theocorp Holding Company (2010)
Study type	Reproductive toxicity - Developmental Toxicity / Teratogenicity (GLP)
Species	Rabbit (New Zealand)
Route	Oral (diet)
Dose level	0, 0.0625, 0.125, or 0.188 % (approximately 0, 21, 41, or 63 mg/kg bw/d)
Purity	>99.5 % purity (confirmed by HPLC/MS, gas-liquid chromatography, and infrared spectroscopy)
Treatment period	GD 6-29
OECD GL	-
Klimisch score	2 (Reliable with restriction)
NOAEL	NOAEL maternal toxicity: 21 mg/kg bw/d NOAEL fetotoxicity/teratogenicity: 21 mg/kg bw/d
Dams	41 and 63 mg/kg bw/d: significant reductions in food intake on GD 15-30
Foetuses	41 and 63 mg/kg bw/d: decrease in foetal body weight, significant increases in the frequency of skeletal variations, indicating a delay in osteogenesis. Effects are considered to be related to the significant reductions in food intake on GD 15-30. Additionally, reduced litter numbers in the control group resulted in increased foetal weights for comparative purposes with treatments. Overall, dietary theobromine was neither teratogenic nor embryotoxic to rabbits in this study.

Paraxanthine

Test substance	Paraxanthine
Reference	York *et al* (1986)
Study type	Reproductive toxicity - Developmental Toxicity / Teratogenicity
Species	Mice (C57BL/6J)
No. per group	16 (control), 13 (low dose), 16 (high dose)
Route	Intraperitoneal (IP)
Dose level	0, 175 or 300 mg/kg bw (dissolved in deionised water)
Treatment period	GD 11-12
OECD GL	-
Klimisch score	2 (Reliable with restriction)
NOAEL	NOAEL maternal toxicity: 175 mg/kg bw NOAEL fetotoxicity/teratogenicity: <175 mg/kg bw
Dams	300 mg/kg bw: maternal toxicity (mortality 1/16, few animals showed convulsions)
Foetuses	175 and 300 mg/kg bw: decreased mean foetal weight (not dose-dependent; control 1.03 g, low dose 0.95 g, high dose 0.94 g), decreased number of implants (not dose-dependent; control 130, low dose 109, high dose 115), increased resorption rate (not dose-dependent; control 12%, low dose 20.2%, high dose 20.9%), dose-dependent increase in malformations (cleft palate, limb; control 0%, low dose 3.4%, high dose 46.2%)

Caffeine

Test substance	Caffeine	
Reference	OECD (2002)	
Study type	Repeated dose toxicity study (90 days)	
Species	Rat (F344/N)	
Route	Oral (drinking water)	
No. per group	12/sex	
Dose level	0, 188, 375, 750, 1500, and 3000 ppm in the drinking water (male: ca. 19.7, 42, 85.4, 151, 272 mg/kg bw/d; female: 23, 51, 104, 174, 287 mg/kg bw/d)	
OECD GL	- (NTP design)	
Klimisch score	2 (Reliable with restriction)	
	Male	*Female*
NOAEL	NOAEL 1500 ppm (151 mg/kg bw/d males)	NOAEL 1500 ppm (174 mg/kg bw/d females)
Mortality	none	none
Body weight	3000 ppm: significant bw reduction (26%)	3000 ppm: significant bw reduction (20%)
Haematology	No relevant findings	No relevant findings
Organ weights	No relevant findings	No relevant findings
Histopathological findings	No relevant findings, dose-dependent cellular enlargement in salivary gland was considered adaptive (reversible effects in the salivary glands are a well-known pharmacological effect of caffeine, sympathomimetic).	No relevant findings, dose-dependent cellular enlargement in salivary gland was considered adaptive (reversible effects in the salivary glands are a well-known pharmacological effect of caffeine, sympathomimetic).

Test substance	Caffeine	
Reference	OECD (2002)	
Study type	Repeated dose toxicity study (90 days)	
Species	Mouse (B6C3F1)	
Route	Oral (drinking water)	
No. per group	12/sex	
Dose level	0, 94, 188, 375, 750, 1500 ppm (males: ca. 21, 44, 85, 130, 167 mg/kg bw/d; females: 25, 47, 88, 134, 180 mg/kg bw/d)	
OECD GL	- (NTP design)	
Klimisch score	2 (Reliable with restriction)	
	Male	*Female*
NOAEL	NOAEL 1500 ppm (167 mg/kg bw/d)	NOAEL 1500 ppm (180 mg/kg bw/d)
Mortality	none	none
Body weight	No relevant findings	No relevant findings
Haematology	No relevant findings	No relevant findings
Organ weights	No relevant findings	No relevant findings
Histopathological findings	No relevant findings, at 1500 ppm adaptive changes in the salivary glands (reversible effects in the salivary glands are a well- known pharmacological effect of caffeine, sympathomimetic).	No relevant findings, at 1500 ppm adaptive changes in the salivary glands (reversible effects in the salivary glands are a well-known pharmacological effect of caffeine, sympathomimetic).

Test substance	Caffeine
Reference	OECD (2002)
Study type	Reproductive toxicity - Developmental Toxicity / Teratogenicity
Species	Rat (Sprague-Dawley CD), 20 per group
Route	Oral (gavage)
Dose level	0, 40, and 80 mg/kg bw/d
Treatment period	Gestation days (GD) 1 to 19
OECD GL	-
Klimisch score	2 (Reliable with restriction)
NOAEL	LOAEL 40 mg/kg bw/d (maternal toxicity), NOAEL 40 mg/kg bw/d (fetotoxicity), NOAEL 80 mg/kg bw/d (teratogenicity)
Dams	40 and 80 mg/kg bw/d: significant decrease in bw gain (35% and 43% respectively)
Foetuses	80 mg/kg bw/d: significant reduction in foetal bw gain (13%)

Test substance	Caffeine
Reference	OECD (2002)
Study type	Reproductive toxicity - Developmental Toxicity / Teratogenicity
Species	Rat (Sprague-Dawley CD), 12 per group
Route	Oral (gavage)
Dose level	0, 10, 20, and 40 mg/kg bw/d
Treatment period	Gestation days (GD) 1 to 20 (dams were allowed to deliver)
OECD GL	-
Klimisch score	2 (Reliable with restriction)
NOAEL	LOAEL 10 mg/kg bw/d (maternal toxicity), NOAEL 40 mg/kg bw/d (fetotoxicity/teratogenicity)
Dams	10, 20, 40 mg/kg bw/d: significant decrease in bw gain (14%, 12% and 18% respectively)
Foetuses	No relevant findings