



Beyond AOPs: A Mechanistic Evaluation of NAMs in DART Testing

Ramya Rajagopal*, Maria T. Baltazar, Paul L. Carmichael, Matthew P. Dent, Julia Head, Hequn Li, Iris Muller, Joe Reynolds, Kritika Sadh, Wendy Simpson, Sandrine Spriggs, Andrew White and Predrag Kukic

Unilever Safety and Environmental Assurance Centre, Colworth Science Park, Sharnbrook, United Kingdom

OPEN ACCESS

Edited by:

Daniel Villeneuve,
United States Environmental
Protection Agency, United States

Reviewed by:

Terje Svengen,
Technical University of Denmark,
Denmark

Olavi R. Pelkonen,
University of Oulu, Finland
Karine Audouze,
Université de Paris, France

*Correspondence:

Ramya Rajagopal
ramya.rajagopal@unilever.com

Specialty section:

This article was submitted to
In Vitro Toxicology,
a section of the journal
Frontiers in Toxicology

Received: 17 December 2021

Accepted: 31 January 2022

Published: 07 March 2022

Citation:

Rajagopal R, Baltazar MT,
Carmichael PL, Dent MP, Head J, Li H,
Muller I, Reynolds J, Sadh K,
Simpson W, Spriggs S, White A and
Kukic P (2022) Beyond AOPs: A
Mechanistic Evaluation of NAMs in
DART Testing.
Front. Toxicology. 4:838466.
doi: 10.3389/ftox.2022.838466

New Approach Methodologies (NAMs) promise to offer a unique opportunity to enable human-relevant safety decisions to be made without the need for animal testing in the context of exposure-driven Next Generation Risk Assessment (NGRA). Protecting human health against the potential effects a chemical may have on embryo-foetal development and/or aspects of reproductive biology using NGRA is particularly challenging. These are not single endpoint or health effects and risk assessments have traditionally relied on data from Developmental and Reproductive Toxicity (DART) tests in animals. There are numerous Adverse Outcome Pathways (AOPs) that can lead to DART, which means defining and developing strict testing strategies for every AOP, to predict apical outcomes, is neither a tenable goal nor a necessity to ensure NAM-based safety assessments are fit-for-purpose. Instead, a pragmatic approach is needed that uses the available knowledge and data to ensure NAM-based exposure-led safety assessments are sufficiently protective. To this end, the mechanistic and biological coverage of existing NAMs for DART were assessed and gaps to be addressed were identified, allowing the development of an approach that relies on generating data relevant to the overall mechanisms involved in human reproduction and embryo-foetal development. Using the knowledge of cellular processes and signalling pathways underlying the key stages in reproduction and development, we have developed a broad outline of endpoints informative of DART. When the existing NAMs were compared against this outline to determine whether they provide comprehensive coverage when integrated in a framework, we found them to generally cover the reproductive and developmental processes underlying the traditionally evaluated apical endpoint studies. The application of this safety assessment framework is illustrated using an exposure-led case study.

Keywords: DART, NAMs, non-animal alternatives, NGRA, mechanistic evaluation

1 INTRODUCTION

The value of using New Approach Methodologies (NAMs) as alternatives to animal testing in the evaluation of chemical safety has gained much attention and recognition (ECC/HC, 2016; ECHA, 2016; ECHA, 2017; EPA, 2018). To foster their development and application, various agencies have put forward guidance, frameworks and workplans that ensure confidence, consistency and are fit-for-purpose, when generating NAMs hazard data for various purposes (OECD, 2005; OECD, 2017; OECD, 2018; Parish et al., 2020). The International Cooperation on Cosmetics Regulation (ICCR) has outlined key principles that guide risk assessors to use NAMs in an integrated manner for Next

Generation Risk Assessment (NGRA) (Dent et al., 2018). There is also additional guidance published on specific NAMs that can be combined in a risk assessment of cosmetic ingredients, in a tiered manner, aligned to the different tiers of the SEURAT-1 *ab initio* workflow for systemic repeat-dose toxicity (Berggren et al., 2017; ICCR, 2018). Case studies are being generated to illustrate the practical application of these principles in assessing the safety of cosmetic products and demonstrating that NAMs can provide valuable insights into non-animal safety assessment (Baltazar et al., 2020).

For many substances, data covering developmental and reproductive toxicity (DART) are lacking, and filling these gaps using traditional methods would use a vast number of experimental animals (Rovida and Hartung, 2009). The development of NAMs to address DART effects is therefore a high priority. Several projects have evaluated the predictive value of batteries of alternative methods for DART, that target a select set of mechanisms and adverse outcome pathways (AOPs). The ReProTect project, which was part of the 6th European Framework Program, published the findings of its feasibility study where 10 chemicals were tested in a set of 14 assays. Based on a nearest neighbour and weight of evidence approach, the results were used to predict any adverse effects on fertility and embryonic development (Schenk et al., 2010). The 7th European Framework Program's ChemScreen project published the DART outcome of testing 12 compounds on a set of 31 assays. Toxicokinetic modelling was included to validate the *in vitro* to *in vivo* dose comparisons and an *in silico* pre-screening module was also applied to reduce testing needs (Piersma et al., 2013; van der Burg et al., 2015). These important studies shared a common goal; to use NAMs to predict whether a substance will be a reproductive or developmental toxicant *in vivo*, often using animal test data as a benchmark. However, in many cases the uptake of these methods into regulatory decision making has been slow, which is understandable given the complexities of trying to use individual NAMs to predict apical endpoints in an intact organism. In recent years, a lot of emphasis has been placed on using AOPs to organise the mechanistic data (including data generated using NAMs) to facilitate safety decision making and either predict or explain adverse outcomes (Ankley et al., 2010; Meek et al., 2014). Eleven AOPs from AOP Wiki, pertaining to DART outcomes, were initially cited by Knapen et al. and since then others have been added for consideration (Knapen et al., 2015). However, many of these to a large extent still remain in development, limited by biological understanding and supporting evidence, and in some cases, a single AOP may not be sufficient to explain an AO and a network approach may be needed to explain or predict the final AO. A few examples are those of human hepatotoxicity, human neurotoxicity, swim bladder inflation in fishes, and male rat reproductive tract abnormalities, each of which are outlined by several distinct AOPs (Villeneuve et al., 2014; Spinu et al., 2019; Arnesdotter et al., 2021). Therefore, while the AOPs are very useful in development of NAMs, such an approach for even more complex outcomes like DART, especially in cases where the MoA of a chemical is unknown, is particularly challenging.

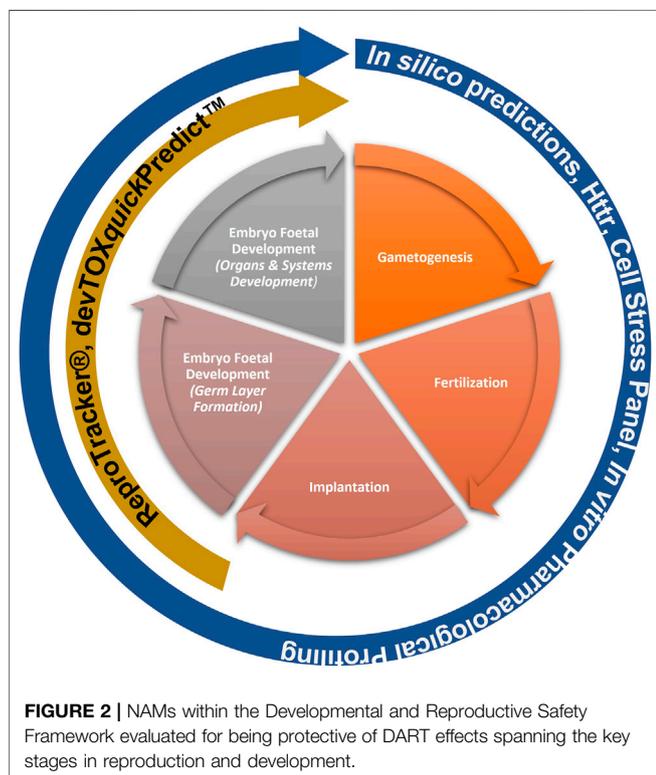
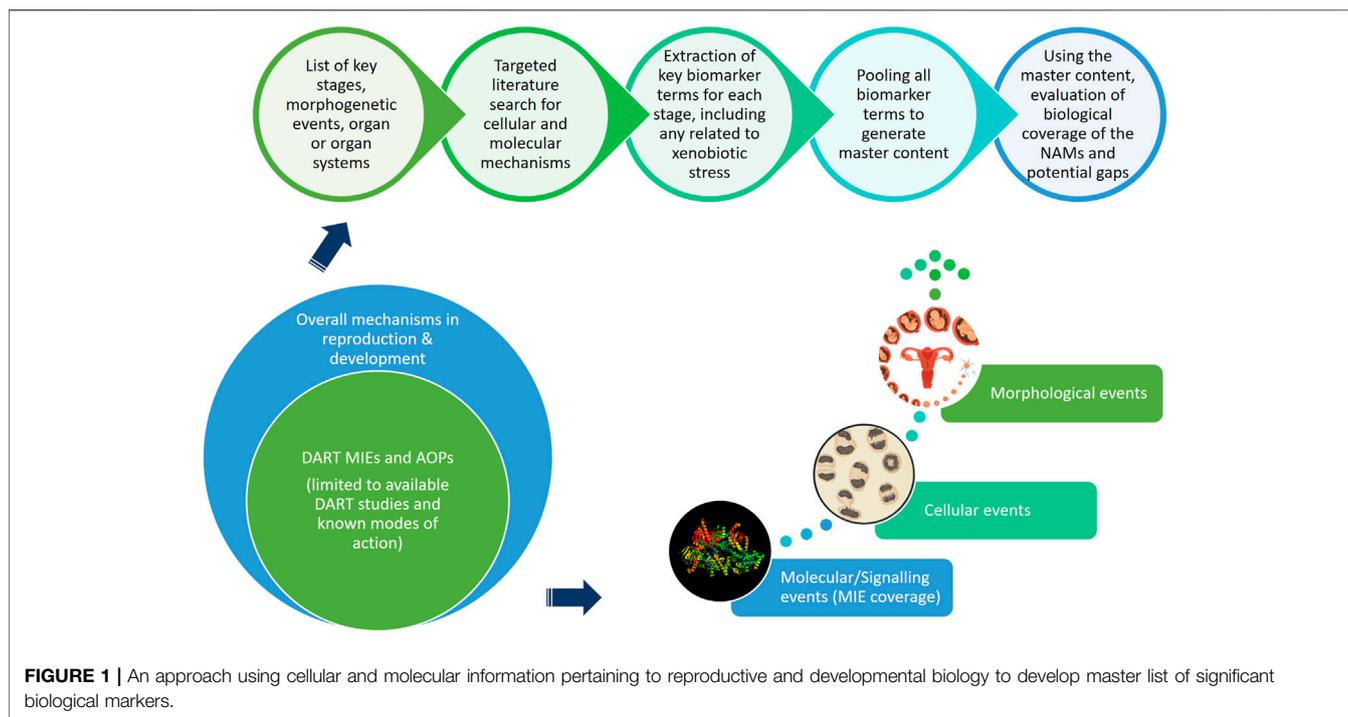
However, it is questionable whether prediction of specific adverse outcomes is necessary to enable decisions to be made on the safe use of chemicals. Do we need to be able to predict adverse DART outcomes (e.g., hypospadias, cleft palate, fused vertebrae), or is it more useful and relevant to know that under specified exposure conditions, an adverse DART outcome is not likely to happen? This change in mindset from a desire to *predict* adverse effects in high dose animal tests to an approach that seeks to *protect* humans from harm at relevant exposures provides the opportunity to allow context-dependent safety decisions to be made using the non-animal tools and approaches available today and to prevent unnecessary animal testing.

This “protection, not prediction” philosophy set the context for the current work, which aimed to take advantage of the knowledge of mechanisms and signalling pathways reported to play a key role in human reproduction and embryo-foetal development, to evaluate how protective a DART NAM framework is for DART effects. As all morphological events or physiological processes are underpinned by cellular events and these are in turn orchestrated by molecular signalling events, gathering this cellular and molecular information pertaining to reproductive and developmental biology is a useful approach in developing a master list of biological markers of significance (Figure 1). A list of key stages and developmental landmarks, morphogenetic events, organ or organ systems was developed, followed by a systematic and targeted literature search conducted for cellular and molecular mechanisms underlying each of these stages. Relevant biological marker terms were then extracted from the abstracts of these publications, which were pooled to generate metadata of markers in human reproduction and development. Additionally, any markers related to general xenobiotic stress pathways and processes were also extracted into this metadata. This master content of biomarkers was used to evaluate the mechanistic coverage of the NAMs within our developmental and reproductive safety framework (Figure 2), assembled from the tools with the most biological coverage in the NGRA framework for systemic toxicity, outlined in Baltazar et al. (Baltazar et al., 2020), and newly available NAMs for detecting developmental toxicity. The overlay of the key set of markers upon the DART framework was also used to identify potential redundancies or gaps to be addressed. Finally, we exemplified our approach with a case study, where DART related NAM hazard data generated for caffeine is assessed against hypothetical exposure scenarios to explore how a framework could aid use of this information in decision making without generating any animal data.

2 MATERIALS AND METHODS

2.1 Literature Search and Review

Guided by the overall knowledge of human reproductive biology and embryo-foetal development, the key stages and morphogenetic events, listed in Table 1, were considered for targeted literature search (Gilbert and Barresi 2016; Gabbe et al., 2017; Kliegman et al., 2020).



The EPA-developed Abstract Sifter literature review tool (<https://github.com/USEPA/CompTox-Chemistry-Dashboard-Abstract-Sifter>) was used to generate a set of relevant literature for each of the key stages listed above (Baker et al., 2017). In order to

maximise the reports that either list or discuss biological markers, either related to signalling pathways or cellular processes, in the context of embryonic development, the query terms were standardised accordingly and aligned with the MeSH terms. Further, the focus of the literature search was on findings in human or mammalian systems and other vertebrate or invertebrate literature were excluded. Any reports related to infections were also excluded, due to limited value in understanding of developmental mechanisms. However, reports related to tumours or cancers were included as they can provide insights into normal developmental mechanisms indirectly. A standard query run was as follows: where x denotes specific terms related to each of the key stages and organs or systems from the master list. Additional inclusion terms related to specific physiology (e.g., nervous system or digestive system) or exclusion terms were used in a bespoke manner. Query terms used for each of the searches are provided in **Supplementary File S1**. To capture the use of different nomenclature describing the same term, index terms from the MeSH ontology were used, wherever possible, that included its synonyms, sub-tree terms and their synonyms. By using index terms, such as “fetal development” the search automatically expands to its synonyms (“fetal growth,” “fetal programming”), all sub-tree terms (“fetal movement,” “fetal organ maturity,” “fetal viability,” “fetal weight,” “gestational age”) and their synonyms.

“(x) AND (embryonic development OR fetal development) AND (cell physiology OR signalling OR pathway OR gene OR protein) AND (human OR mammalian) NOT (amphibians OR frog) NOT (fishes)

TABLE 1 | Key stages, morphogenetics events and derivatives of human reproductive biology and embryo-foetal development.

Stage no	Key stages and morphogenetic events	Derivative organs and systems
1	Sex determination	
2	Gametogenesis	
3	Fertilization	
4	Zygote formation	
5	Implantation	
6	Blastulation	
7	Gastrulation	
8	Placenta formation	
9	Neurulation	
10	Ectoderm formation and its derivatives	a. Central nervous system b. Peripheral nervous system c. Autonomus nervous system d. Integumentary system
11	Mesoderm formation and its derivatives	a. Somitogenesis b. Hematopoiesis c. Heart and circulatory system d. Immune system e. Spleen f. Urinary system and urethra g. Reproductive system—testis h. Reproductive system—ovary i. Skeletal system j. Limbs
12	Endoderm formation and its derivatives	a. Digestive system b. Respiratory system c. Thymus d. Parathyroid e. Thyroid
13	Structures developing from mesenchyme or multiple germ layers	a. Adrenal glands b. Eyes c. Ears d. Face and neck
14	Intrauterine growth	

NOT (invertebrates) NOT (birds OR aves OR avian OR chick) NOT (infections)”

Each set of the Abstract Sifter search results were saved using a filename indicating the respective stage or organ or system (**Supplementary File S1**). A quality check was performed manually to ensure the reports were relevant to the search criteria and minimise the number of unrelated results in cases when same terminology is used to describe different phenomena. This quality check review was conducted using the sifter terms functionality, where the key terms from the biology of interest were used (e.g., for central nervous system, the sifter terms included central nervous, spinal, neuron, neural, glial, astrocyte, etc.). Any report(s) that did not contain a single relevant key term but could not be excluded cleanly using specific exclusion terms were further highlighted for exclusion in the subsequent step of biomarker terms extraction.

2.2 Extraction of Key Biomarker Terms

Extraction of key biomarker terms was done using articles published in PubMed on or before 1 August 2021. For each stage/morphological event, related abstracts were first retrieved using relevant PubMed IDs (PMIDs) and then collated in a single

text file. TERMite (<https://www.scibite.com/platform/termite/>), SciBite's named entity recognition engine, was used to align unstructured text to extensively curated vocabularies, during a process of semantic enrichment. Semantic enrichment was limited to three vocabularies, namely, GENEBOOST, miRNA and DrBP (DART-related Biological Processes).

GENEBOOST is SciBite's vocabulary based on the Hugo Gene Nomenclature Committee standard list of genes that covers all human protein coding genes (Tweedie et al., 2021; DOI:10.1093/nar/gkaa980).

miRNA is SciBite's algorithmic module that detects miRNA terms present in the text.

DrBP is a bespoke vocabulary that was generated for the purpose of this work. The vocabulary includes relevant cellular and molecular mechanisms that underly each of the stages pertaining to reproductive and developmental biology. Specifically, the vocabulary was formed by merging terms relevant to signalling pathways and cellular processes contained in the publicly available Gene Ontology (GO) and National Cancer Institute Thesaurus (NCIT) (Ashburner et al., 2000) (<https://ncit.nci.nih.gov/ncitbrowser/>). Namely, it contains: 1) Cell Stress Process class from NCIT (id: C21065) that includes cellular or subcellular processes involved in disturbance or

restoration of a homeostatic condition; 2) Epigenetic Process class from NCIT (id: C21051) that includes terms related to changes in the regulation of the expression of gene activity without alteration of genetic structure; 3) Mitochondrial Damage class from NCIT (id: C45524) that includes any process that leads to dysfunction of mitochondria, whether by oxidative damage, mutation of mitochondrial DNA or other by means; 4) Cellular Process class from GO (id: 0009987) that includes any process that is carried out at the cellular level, but not necessarily restricted to a single cell; 5) Signalling class from GO (id: 0023052) that includes processes in which information is transmitted within a biological system.

For the terms extracted using GENEBOOST and DrBP alone, a threshold of ≥ 10 and ≥ 5 hit counts, respectively, was decided to maximise the relevance of the extracted key biomarker terms. For the terms extracted using the miRNA module, this threshold was set to at least two appearances of a relevant miRNA term in all abstracts associated with the specific key stage/morphological event from **Table 1**.

3 BIOLOGICAL COVERAGE OF NAMs

3.1 Baseline Gene Expression of MCF7, HepG2 and HepaRG Cells in High Throughput Transcriptomics (HTTR)

The 3 cell lines that are currently core to our systemic toolbox are MCF-7, HepG2 and HepaRG cells. To enable a comparison between the key biomarkers identified above using GENEBOOST and the coverage provided by these cell lines, baseline gene expression for each cell line was obtained by analysing gene counts for 24 h with 0.5% DMSO as a vehicle control. The data analysis pipeline followed Reynolds et al. (Reynolds et al., 2020). According to Reynolds et al., probes with a mean or median count less than five across all non-control treatments were discarded. Applying this procedure, we have created three separate gene lists that establish a baseline expression for each of the cell line used in the experiments (**Supplementary File S2**).

3.2 Baseline Gene Expression of iPSCs

The Stemina DevToxQuickPredict and Toxys ReproTracker assays use human induced pluripotent stem cells (iPSCs). An example of baseline gene expression of iPSCs in undifferentiated state was obtained by analysing previously deposited gene expression of eight human iPSCs from StemCellDB (Mallon et al., 2013). The dataset represents quantile normalized log (base = 2) gene expression data from Agilent one-color gene expression microarrays. According to Mallon et al., a gene was considered expressed in this data set if it was detected in at least one of the eight iPSCs using a noise cut-off of 7.5. Applying this procedure, we have created a gene list which defines pluripotency and establishes a baseline for the undifferentiated pluripotent state (Mallon et al., 2013) (**Supplementary File S3**), which was then utilized to compare with the key biomarkers extracted with GENEBOOST.

3.3 *In Vitro* Pharmacological Profiling (IPP)

The *In vitro* Pharmacological Profiling (IPP) panel contains 72 binding, enzymatic and coactivator recruitment assays associated with 60 targets with known safety liabilities. 44 of the targets have been previously associated with *in vivo* adverse drug reactions (Bowes et al., 2012). An additional 16 targets implicated in developmental toxicity were added to the panel based on available literature (National Research Council et al., 2000; Wu et al., 2013). They include six nuclear hormone receptors (Estrogen Receptor Alpha, Estrogen Receptor Beta, Thyroid Hormone Receptor Alpha, Thyroid Hormone Receptor Beta, Progesterone Receptor and Mineralocorticoid Receptor) and 10 basic helix-loop-helix transcription factors (Retinoic acid receptor alpha, Retinoic acid receptor gamma, Retinoic acid receptor RXR-alpha, Retinoic acid receptor RXR-beta, Peroxisome proliferator-activated receptor alpha, Peroxisome proliferator-activated receptor delta, Peroxisome proliferator-activated receptor gamma, Pregnane X receptor, Constitutive androstane receptor, Vitamin D Receptor). The full list of the IPP targets (**Supplementary File S4**) was compared with the list of biomarkers identified using DrBP.

3.4 Cell Stress Panel (CSP)

Cell stress panel comprises biomarkers that cover eight key stress pathways (oxidative stress, inflammation, ER stress, metal stress, DNA damage, heat shock response, hypoxia and translocation of Aryl Hydrocarbon Receptor), mitochondrial toxicity and general cell health (Simmons et al., 2009). For more information on the specific list of assays included in the panel refer to (Hatherell et al., 2020). The CSP assays were compared with the list of biomarkers identified using DrBP.

3.5 devTOXquickPredict™

devTOXquickPredict™ is a targeted exposure-based biomarker assay on human embryonic stem cells (hESCs) or induced pluripotent cells (iPSCs), where in response to a nine-point dose response testing, specific metabolites, along with a cytotoxicity endpoint, are monitored. Specifically, the ratio of ornithine to cystine is measured as they were identified as indicators of developmental toxicity (Palmer et al., 2013; Palmer et al., 2017). This output was used to assess the coverage of certain specific cellular processes identified by DrBP.

3.6 ReproTracker®

ReproTracker® assesses any perturbation, following chemical exposure, to the process of differentiation of stem cells to specific germ layers and the derivatives therein, by monitoring specific biomarkers (Racz et al., 2018) (Jamalpoor et al., submitted 2021). Human induced pluripotent stem cells (hiPSCs), differentiating into endoderm, mesoderm or ectoderm cell types and further to derivative hepatocytes, cardiomyocytes and neural rosettes, respectively, were treated with six different doses of the test chemical. Any change in levels of AFP, MYH6 and Pax6 expression as well as other morphological features were monitored, as markers for the differentiation of liver, heart and neural cells, respectively, and thereby potential teratogenic effects were determined. These read-outs were compared with the key

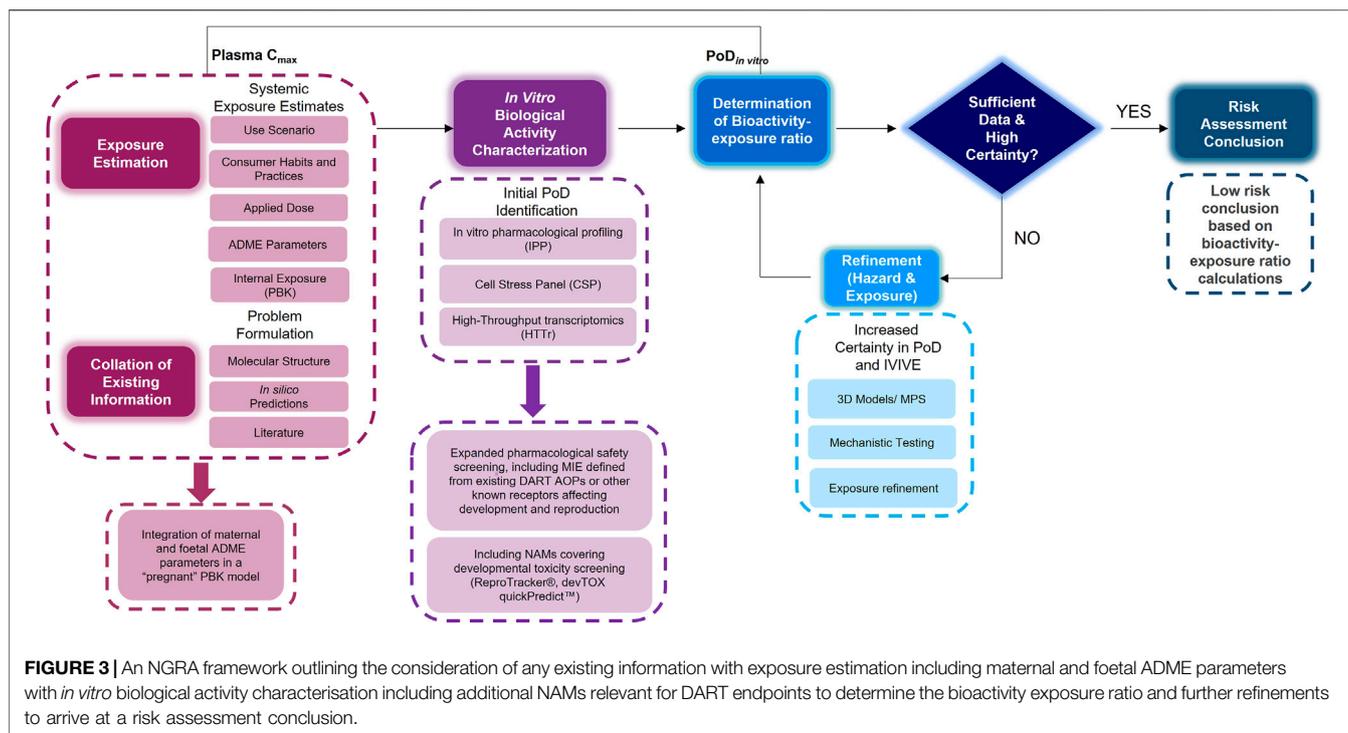


FIGURE 3 | An NGRA framework outlining the consideration of any existing information with exposure estimation including maternal and foetal ADME parameters with *in vitro* biological activity characterisation including additional NAMs relevant for DART endpoints to determine the bioactivity exposure ratio and further refinements to arrive at a risk assessment conclusion.

biomarkers extracted by DrBP, specifically those related to differentiation processes.

3.7 *In Silico* Predictions

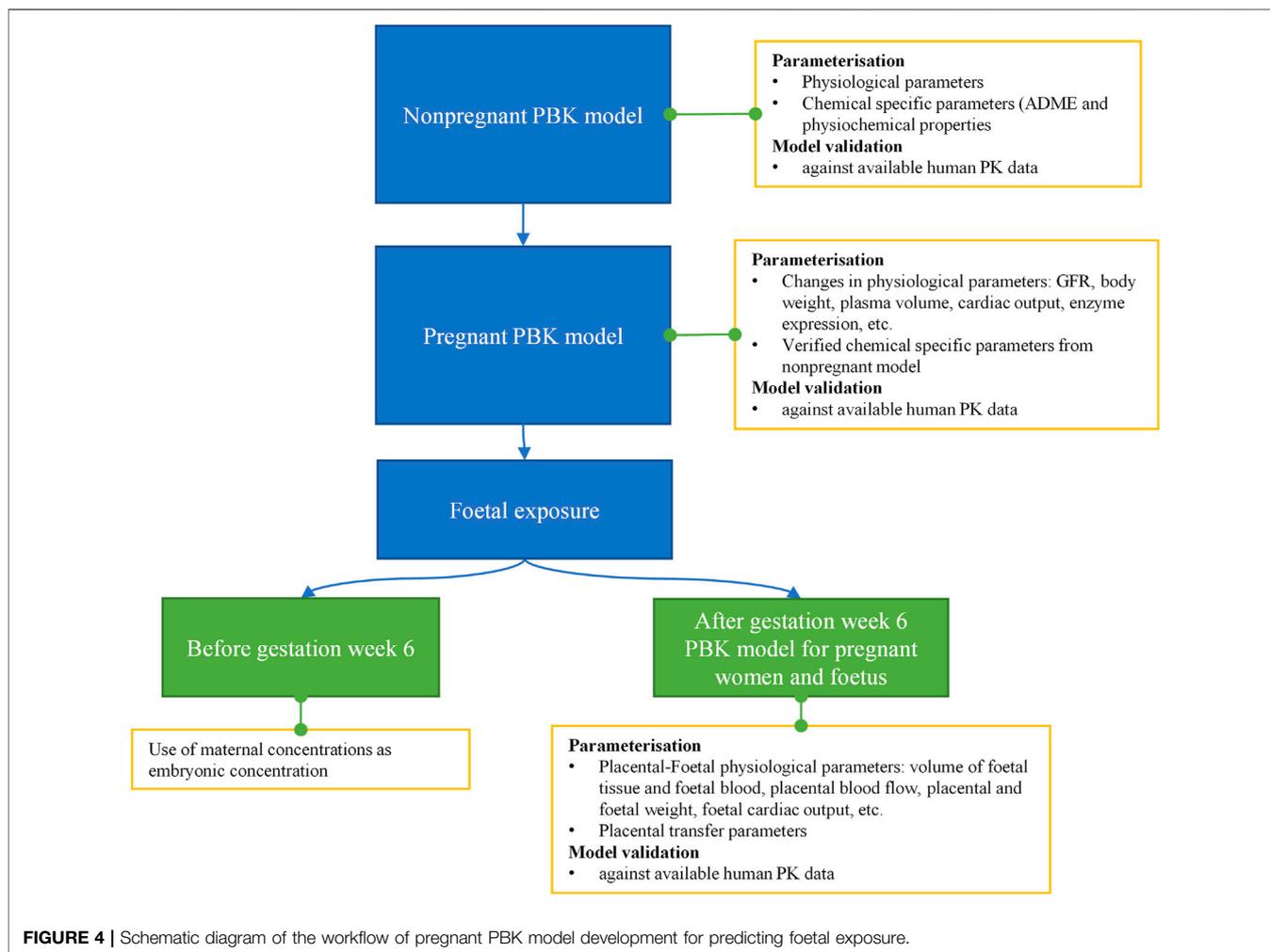
Based on accurate representations of chemical structure, *in silico* predictions can be used to inform on the potential biological activity profile of the molecule, give indication of any possible selective or non-selective activity and provide a broad toxicity screen across many apical endpoints based on the existing toxicity data. These predictions were used as weight of evidence for biological coverage. For more information on the list of *in silico* tools used as a part of the DART Toolbox and their associated endpoint refer to **Supplementary Table S3**.

4 DEVELOPMENT OF A NGRA DART FRAMEWORK AND APPLICATION TO THE RISK ASSESSMENT OF CAFFEINE IN CONSUMER PRODUCTS

An NGRA workflow, illustrating how to integrate NAMs for the systemic safety assessment of coumarin has been previously published (Baltazar et al., 2020). Despite the workflow being tailored to the coumarin case study, its key building blocks are applicable to any NGRA. We hypothesize that the NAMs described in **Figure 2** could form a core toolbox within an NGRA framework which is protective of DART effects (**Figure 3**). If there is sufficient confidence that this toolbox provides coverage for the processes that are critical for normal reproduction and development, the bioactivity-exposure ratio

(BER) (ratio between PoD and exposure) can be calculated and, together with all the other available evidence a decision can be made to whether there is sufficient information and high certainty to reach a risk assessment conclusion. If the outcome is uncertain, an iterative process can begin by designing new experiments addressing specific gaps (**Figure 3**).

Caffeine was selected as a case study chemical to exemplify the use of the DART framework, based on the availability of human exposure data from cosmetics and dietary sources and known mode of action. High consumption of caffeine (>200 mg/day) during pregnancy has been associated with increased risk of adverse birth weight-related outcomes (i.e., foetal growth retardation, small for gestational age), and therefore the European Food Safety Authority (EFSA) has advised a limit of 200 mg/day for pregnant women instead of the 400 mg/day for the general adult population (EFSA, 2015). The objective of this case study was to investigate how protective our NGRA DART framework is for chemicals suspected of causing reproductive and developmental adverse health effects in humans. To exemplify how this can be achieved, two distinct exposure scenarios were selected, a dermal exposure to 0.1% caffeine in a hypothetical body lotion and an oral consumption of 200 mg/day of caffeine at different gestational stages of pregnancy (6, 20 and 30 weeks). After determining exposure, *in silico* tools (see *In Silico* Predictions) were run followed by the generation of the *in vitro* biological activity data, including the IPP, CSP, HTTr, ReproTracker® and devTOXquickPredict™. PoDs derived from the CSP, HTTr and IPP were compared to exposure estimates (maternal and foetal plasma C_{max}) to calculate a bioactivity-exposure ratio (BER).



4.1 Exposure Estimation: PBPK Modelling

4.1.1 Workflow

A PBK model was built to make predictions on both maternal and foetal concentration (i.e., plasma C_{max}) of caffeine under two exposure scenarios, i.e., oral administration of 200 mg caffeine and topical application of a hypothetical body lotion containing 0.1% caffeine, so that the C_{max} concentrations could be compared against *in vitro* PoD values to derive the BER. Although foetal concentration is a more relevant dose metric for a substance suspected of being a direct developmental toxicant, both maternal and foetal concentrations were predicted throughout different trimesters in pregnancy, and C_{max} from all the predicted maternal and foetal concentrations were used in the BER calculation to be conservative in decision making.

The workflow we used to develop the PBK model is shown in **Figure 4**. Firstly, a non-pregnant PBK model was built and verified against available observed human clinical data from both oral and dermal exposures routes. In case the predicted PK profile and parameters deviated from the observed data, the model was refined by parameter optimization through fitting against the non-pregnant clinical data. Secondly, the pregnancy

PBK model was developed with the verified chemical-specific parameters and pregnancy-related physiological changes over time, and then further validated against available human PK data in pregnancy. Thirdly, the verified pregnancy PBK model was extended for embryonic/foetal concentration prediction at different gestational ages. In the foetal period of development, foetal circulation is established during the early stages (week 5–6 of gestation), allowing the growing fetus to receive the required oxygen and nutrients as well as dispose of waste products. This type of circulation involves the umbilical cord and placental blood vessels which carry foetal blood between the foetus and the placenta (Wang and Zhao, 2010). Therefore, in the pregnancy PBK model, week 6 of gestational age (GA) was chosen to separate two stages. For GA before week 6, only the uterus tissue is added to the model; for GA above week 6, in addition to the uterus tissue, placenta tissues, amniotic fluid, foetal tissue and foetal blood circulation are also included in the model. Lastly, the extended pregnancy PBK model was used to make predictions on plasma concentrations in both mother and embryo/foetus at different GAs under the two exposure scenarios of interest.

4.1.2 Software

PBK models for caffeine were built using the software GastroPlus 9.8 (Simulation Plus, Lancaster, CA, United States). Population-dependent physiological parameters for non-pregnant and pregnant PBK models were obtained using the Population Estimates for Age-Related (PEAR) Physiology module in GastroPlus. Pregnancy incurred changes in the maternal body: weight gain, changes in enzyme expression levels, enlargement of certain tissues such as uterus, placenta, brain and kidney, increased amount of plasma volume, GI changes (increased stomach transit time), etc., as well as the volume of foetal tissue and foetal blood with GA have all been incorporated in the built-in pregnancy model. When maternal age, GA and foetal gender are specified, the default physiology will be generated by built-in equations. Alternatively, the default physiologies for both maternal and foetal subjects, such as body weight and height, cardiac output, weight gain and individual tissue weights and perfusions could be manually modified.

4.1.3 Development and Validation of the Non-Pregnant and Pregnant PBK Model

The physiochemical and ADME parameters used to build the PBK models were derived from published *in silico*, *in vitro*, and human PK data, and are presented in **Supplementary Table S1**. Tissue-to-plasma partitioning coefficients were calculated in Gastro-Plus using the Lukacova (Rodgers–Single) method, assuming chemical distribution into the tissues being perfusion-limited. The liver and kidney were considered to be the only organs to eliminate caffeine. Caffeine is mainly metabolized by CYP1A2 (Grzegorzewski et al., 2021). The liver metabolism was described in the model by using *in vitro* K_m and V_{max} values for CYP1A2 reported in the literature, and that the V_{max} was then optimized by GastroPlus to better fit caffeine plasma profiles and predict clearance in non-pregnant clinical data of intravenous infusion (Blanchard and Sawers 1983; Cheng et al., 1990; Otberg et al., 2008). The model was then validated by comparing the simulated PK data with the observed clinical studies from both oral and dermal applications (Blanchard and Sawers 1983; Cheng et al., 1990; Otberg et al., 2008). After the non-pregnant PBK model was developed and verified, the pregnancy PBK model was built with the same organs mentioned in non-pregnant subjects, the verified chemical-specific parameters and pregnancy incurred physiological changes (the physiological changes have been incorporated in GastroPlus for the pregnancy population). It is worth noting that the decrease in CYP1A2 activity for each pregnancy trimester was captured in the model too. Then this pregnancy PBK model was validated by comparing the simulated PK data with the available observed clinical study (Brazier et al., 1983).

4.1.4 Extending the Pregnant PBK Model for Embryonic/Foetal Exposure Prediction

The transfer of substances between the maternal to foetal blood-stream comprises of materno and foetosyncytiotrophoblast exchanges at the apical and basal face of the syncytiotrophoblast, respectively (Codaccioni et al., 2019).

Compounds transfer across the placental *via* passive diffusion, active transport, facilitated diffusion, pinocytosis and phagocytosis. Passive diffusion and active transport, to a lesser extent, are the predominant mechanisms of placental transfer for small molecules, like caffeine (Griffiths and Campbell 2014). In this study, passive diffusion was considered to be the main mechanism for placental transfer of caffeine, as no evidence has been shown that it undergoes active uptake/efflux in placenta. Unlike other tissues, placenta was set to be permeability limited tissue in the pregnancy model. The placental transfer input, i.e., apparent permeability coefficient (P_e) was obtained from *in vitro* placental transfer assays using BeWo b30 cells as reported in literature (**Supplementary Table S1**). This parameter was needed to calculate the PStc (permeability cellular surface area product) in the placental compartment. The cellular surface areas of placenta at different GA were obtained as reported in literature (Griffiths and Campbell, 2014; Dhyani, 2021).

4.1.5 PBK Predictions on Maternal and Foetal Concentrations of Caffeine Under Two Exposure Scenarios

The extended pregnancy PBK model was used to make predictions on plasma concentrations in both mother and embryo/foetus at different GAs under the two exposure scenarios of interest. The simulation of each scenario was run to predict steady state concentration profiles on 30 years old pregnant women. For the topical application, consumer exposure to body lotion was characterised based on SCCS guidance as presented in **Supplementary Table S2** and used in the dermal module in GastroPlus as product use input.

4.2 Collation of Existing Information

Caffeine (CAS: 58-08-2) is a member of the methylxanthine group, commonly found in foods and drink. The suite of *in silico* tools described above was used to provide predictions of toxicity based on caffeine's structure (**Supplementary Table S3**). For completeness, tools were run for non-DART toxicological endpoints in addition to the DART relevant endpoints.

4.3 *In Vitro* Biological Activity Characterisation

In vitro biological activity characterisation was performed by generating data for caffeine in the following tools: CSP, HTTr, IPP, ReproTracker[®] and devTOXquickPredict[™] and methods are summarised below.

Caffeine was tested in cell stress panel according to Hatherell et al. (Hatherell et al., 2020). Concentration-response datasets were analysed with a Bayesian statistical model which infers a distribution for the PoD. The cumulative density function of the distribution, evaluated at a chosen concentration, is regarded as a measure of confidence in perturbation of the measured biomarker at the chosen concentration. Confidence at the maximum tested concentration was reported as the *concentration-dependency score (CDS)*.

Caffeine was tested in multiple cell models in two different experiments using high-throughput transcriptomics TempO-Seq technology (Yeakley et al., 2017; Baltazar et al., 2020). Experiment 1 included MCF7, HepG2 and HepaRG 2D cells and in experiment 2 HepaRG 3D model was included as well as a repeat in HepaRG 2D cells. Cells were cultured following procedure from Baltazar et al. Concentrations were chosen after an initial cytotoxicity study to exclude cytotoxic doses (if present) and the dose range chosen to encompass the two exposure scenarios (0.01, 0.1, 1, 10, 100, and 1,000 μM). Cells were treated for 24 h using 0.5% DMSO as a vehicle control. Three biological replicas were generated for the experiment 1, whereas five biological replicas were generated for the experiment 2. Gene sequencing was performed on cell lysates using the Biospyder Human Whole Transcriptome Assay panel version 1 as previously described by (Yeakley et al., 2017). The data processing pipeline and concentration–response analysis followed a modification of the method developed by Reynolds et al (Reynolds et al., 2020). Briefly, a Bayesian statistical model similar to that developed in (Baltazar et al., 2020) was constructed to derive PoDs for individual probes. Modifications of the analysis method include the use of a log Student's t-Poisson sampling distribution rather than a negative binomial to allow for increased kurtosis in the distribution of counts, thereby reducing sensitivity to outliers. In addition, variance across biological replicates was no longer used to determine the PoD to increase sensitivity. Lastly, the global PoD which is an estimate of a minimum effect concentration across all probes [see (Reynolds et al., 2020) for further details], is calculated from quantiles in the range 25–75 rather than 1–99 to increase stability of the estimate.

IPP experiments were carried out at Eurofins Cerep SA. At first, the screening was done using a fixed concentration of caffeine at 10 μM in two replicates. Compound binding in binding assays was calculated as a percentage inhibition of the binding of a radioactively labelled ligand specific for each target. Compound enzyme inhibition effect was calculated as a percentage inhibition of control enzyme activity. In order to determine IC₅₀ (concentration causing a half-maximal inhibition of the control response) as a measure of potency for the Adenosine 2A receptor inhibition, the assay was performed in a dose response manner at eight concentrations in two replicas. The IC₅₀ value was determined by the Bayesian probabilistic model of the concentration–response curve (Johnstone et al., 2017). The priors for IC₅₀ were set to the median experimental dose, the slope was set to 1.0 and low and high dose responses were set to 0% and 100%, respectively.

Results for caffeine exposure in the devTOXquickPredict™ assays were extracted from Zurlinden et al. In short, H9 cells were treated with up to 500 μM caffeine for 72 h, with media replacement every 24 h. The cell-conditioned media from the final 24 h treatment period was collected for analysis of the targeted biomarker, ornithine and cysteine ratio, and cell viability (Zurlinden et al., 2020). For the ReproTracker®, hiPSCs were differentiated into hepatocytes, cardiomyocytes and neural rosettes, respectively, in the presence of six doses of caffeine (up to 100 μM), for 21, 14 and 13 days, respectively,

and any change in levels of AFP, MYH6 and Pax6 expression as well as other morphological features were monitored.

4.4 Determination of Bioactivity Exposure Ratio (BER)

For a given PoD, the BER was defined as the ratio between the nominal concentration at which the PoD is defined, and the relevant plasma C_{max} estimate from the PBPK modelling for both exposure scenarios. PoDs were derived from the HTr, Cell Stress and iPP panel.

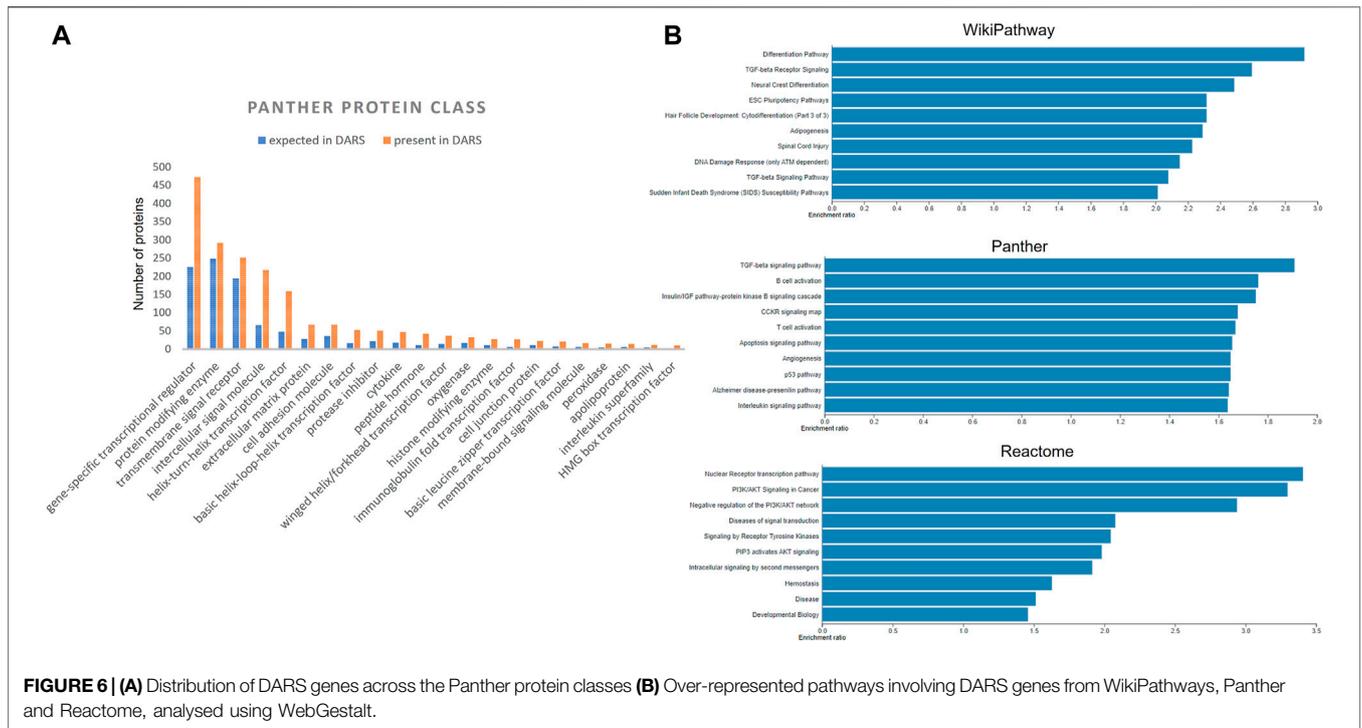
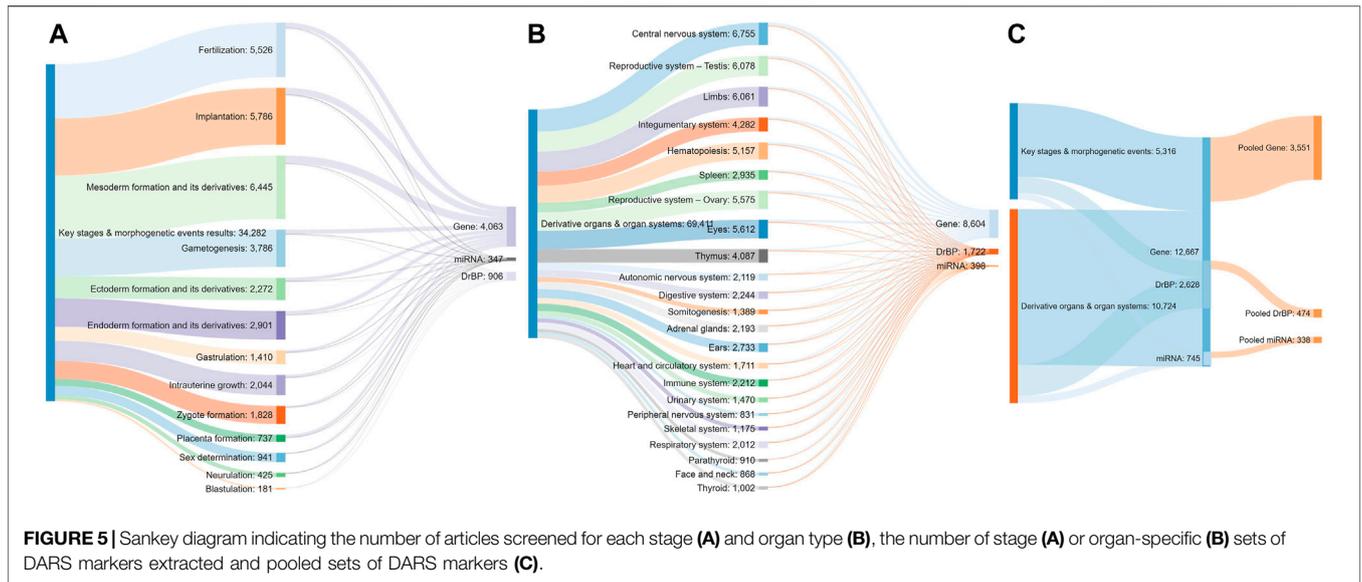
5 RESULTS

5.1 Literature Search and DARS Markers Extraction

A total of 34,308 articles from the literature search on key stages in reproductive biology and embryo–foetal development as well as morphogenetic events (Table 1) were found to be relevant, after validating the results through sifter criteria. Similarly, a total of 69,299 relevant articles were short listed from the searches related to organs and organ systems development (Table 1). This complete list of articles is provided in Supplementary File S1. A certain degree of overlap in articles related to each of the germ layers and their derivatives was seen. However, all the relevant articles were considered for the next steps as an over representation of information was preferred rather than an under representation. The pooling of extracted markers in the subsequent steps addressed any duplications and redundant information. These 103,607 articles served as the comprehensive pool from which biological marker terms relevant to signalling during reproduction and development, referred to as Developmental and Reproductive Signalling (DARS) markers, were extracted (Figure 5). The overall process is depicted in Figure 1.

A named entity recognition engine Termite (SciBite) was used to align the titles and abstracts of the pooled articles to extensively curated vocabularies that represent human genes, cellular and signalling processes, miRNAs and their synonyms (see Methods for more information). This enabled recognition and annotation of multiple terms that are used to describe the same concept (e.g., gene aliases, alternative pathway names, etc). A quality check was performed to ensure the extracted terms are indeed relevant and do not refer to any other meaning. The final list of DARS markers that refer to genes, cellular/signalling processes and miRNAs extracted for each key stage is given in Supplementary Files S5–S7, respectively.

Extracted DARS genes are presented separately for each stage in the development (Supplementary File S5). For instance, they included 902 DARS genes associated with the development of the central nervous system or 785 DARS genes associated with the mesoderm formation. Furthermore, additional validation was carried out by checking if DARS gene sets extracted for each stage were indeed related to the stage in question. This was done by annotating extracted DARS gene sets using the GO biological processes and Human Phenotype Ontology. Then, for each key



stage top 10 statistically overrepresented GO biological processes and phenotypic abnormalities were reviewed and their relevance to the key stage was confirmed (Supplementary File S5). For example, in the gene set associated with the development of the central nervous system, processes and abnormalities such as neuron migration, regulation of neurotransmitter levels, *in utero* embryonic development, neoplasm of central nervous system, neoplasm of the nervous system and morphological abnormalities of the central nervous system were among the

top overrepresented, which supports the conclusion that the associated DARS gene set is indeed very relevant. Lastly, gene sets for each stage in the development were pooled together in a final list of 3,551 DARS genes (Figure 5, Supplementary File S5). Three genes with the highest number of occurrences in the literature were found to be glycoprotein hormones alpha chain (11,924 hits), sonic hedgehog protein (6,622 hits), and proto-oncogene Wnt-1 (6,428 hits). Glycoprotein hormones alpha chain encodes for the alpha chain of the active heterodimeric

glycoprotein hormones, such as thyroid stimulating hormone, luteinizing hormone, follicle stimulating hormone and choriogonadotropin. Sonic hedgehog and Wnt-1 are well-known signalling proteins that play essential role in developmental processes. When the 3,551 DARS genes were classified using protein classes from Panther (Mi et al., 2019), gene-specific transcription regulators, protein modifying enzymes, transmembrane and intercellular signal molecules came up as top four overrepresented classes (**Figure 6A**). In addition, the top overrepresented pathways, analysed using WebGestalt server (Zhang et al., 2005), from WikiPathways (Martens et al., 2020), Panther (Mi et al., 2019) and Reactome (Jassal et al., 2020) were found to be the differentiation pathway, TGF-beta signalling, and nuclear receptor transcription pathway, respectively. For the list of top 10 overrepresented pathways from the three databases refer to **Figure 6B**.

DARS cellular and signalling processes mined in the literature are presented separately for each stage in the development (**Supplementary File S6**). For instance, they included 145 DARS cellular and signalling processes associated with the development of central nervous system or 114 processes associated with the mesoderm formation. When individual processes for each stage in the development were pooled together, they gave a final list of 474 DARS processes (**Figure 5, Supplementary File S6**). Among them, signalling was by far the most mentioned term in the literature with 21,733 occurrences. This was followed by cell cycle and cell death with 3,228 and 2,514 hits, respectively. Other top 10 processes included DNA methylation (2,440 hits), epithelial to mesenchymal transition (2,422 hits), phosphorylation (2,372 hits), cell differentiation (2,262 hits), cell development (2,248 hits) and oocyte maturation (1,973 hits); all of them known to be cardinal in reproductive and developmental biology.

Lastly, DARS miRNAs mined in the literature are presented separately for each stage in the development (**Supplementary File S7**). The number of miRNAs associated with different stages showed high variability from 90 miRNAs associated with the implantation process to no miRNAs detected for the blastulation and development of parathyroid and urethra (see **Supplementary File S7** for all key stages). Furthermore, additional validation was attempted by checking if targets associated with the miRNAs sets extracted for each stage were indeed related to the stage in question by using miRNA-target annotation available in the IPA software (Kramer et al., 2014). However, we were not able to confirm any overrepresented GO biological processes or phenotypic abnormalities that would confirm their relation to different stages in the development. MiRNAs have previously been shown to have diverse roles in fundamental biological processes such as cell proliferation, differentiation, apoptosis and stress response and are important regulators in development [reviewed in (Shenoy and Belloch, 2014; Ivey and Srivastava, 2015)]. While the research on miRNA is steadily increasing, the annotation of miRNAs itself as well as to its biological function is only emerging (Huntley et al., 2018; Kozomara et al., 2019; Alles et al., 2019) which might explain the unsuccessful search in IPA. However, we still believe it is useful to report here the list of DARS miRNAs found to be associated with

the development from the literature. When individual miRNA sets for each stage in the development were pooled together, they gave a final list of 338 DARS miRNAs (**Figure 4, Supplementary File S7**). Top three miRNAs were the let-7 family, miR-21 and miR-145 with 155, 127, and 85 occurrences in the literature, respectively. They have all been previously reported to play an important role in reproductive and developmental biology; the let-7 family is well-known regulator of iPSC reprogramming that promotes differentiation, miR-21 plays a role in cell proliferation, migration, invasion and apoptosis, whereas miR-145 was found to regulate differentiation of multipotent neural crest cells (Ji et al., 2007; Worringer et al., 2014; Sekar et al., 2019). As miRNAs regulate mRNA levels and they form a set of complementary biomarker data to that of gene expression, detailed correlation between the two in a spatial and temporal context would be required to outline their coverage and was decided to be out of scope for this work. For this piece of work, we have focussed on the coverage of transcriptomic markers in greater detail.

5.2 Biological Coverage of DARS Markers

Building on the previously published framework for evaluating systemic toxicity in a safety assessment, which included read-outs from *in silico* predictions, HTTr, cell stress panel and IPP (Baltazar et al., 2020), we developed an NGRA DART framework which includes two additional NAMs relevant for DART endpoints, namely, ReproTracker[®] (Racz et al., 2018) (Jamalpoor et al., submitted 2021) and devTOXquickPredict[™] (Palmer et al., 2013) (**Figure 2**). In order to assess the coverage of human biology and mechanistic relevance of our NGRA DART framework, we set out to evaluate the coverage of individual NAMs integrated in this framework against the identified DARS markers.

5.2.1 Genes

The NGRA framework routinely employs MCF7, HepG2 and HepaRG cell lines for HTTr testing and caffeine was also tested in these 3 cell lines. Hence, as an example of biological coverage of the HTTr assay we identified separate gene lists that established a baseline expression for each of the cell line used in the experiments (see Methods and **Supplementary File S2**). They include 8,931 genes expressed in MCF7, 9,261 genes expressed in HepG2s and 10,819 genes expressed in HepaRG cells. In addition, an example of baseline gene expression of human iPSCs in undifferentiated state was obtained by analysing previously deposited gene expression in eight human iPSCs from StemCellDB (Mallon et al., 2013). Specifically, they include 11,483 genes with details presented in the **Supplementary File S2**. This analysis is used to approximate and inform on the coverage of biology of the devTOXquickPredict[™] assay, which also uses iPSCs in the undifferentiated state. Genes detected in MCF7, HepG2, HepaRG, iPSC cells were pooled together, and their overlap is depicted in **Figure 7**. They include 14,225 genes in total. Almost half of those genes (6,564 genes) were found to be expressed in all four cell lines. Among the cell lines, iPSCs express 2,319 genes that are not present in any of the remaining cell lines referred to as unique genes here; HepaRGs, MCF7s and HepG2s express 781, 186 and 145 unique genes, respectively (**Figure 7A**).

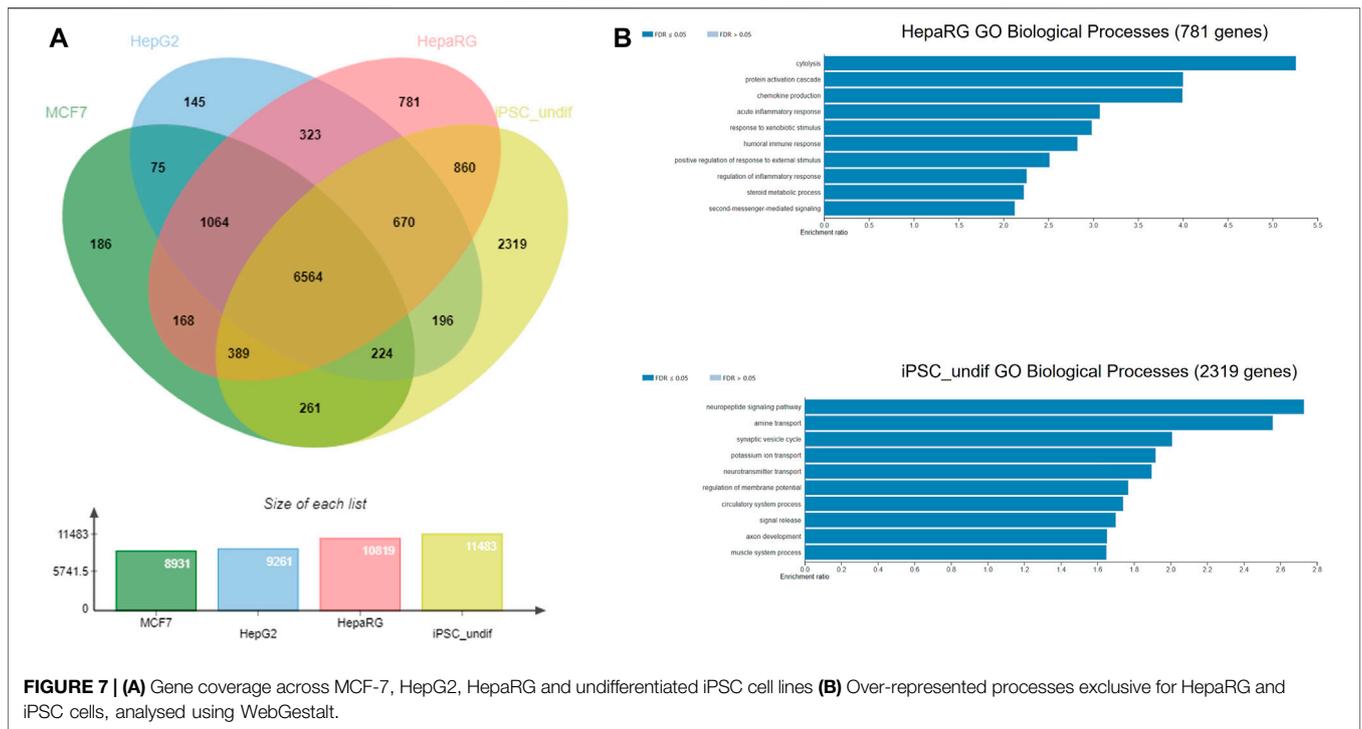


FIGURE 7 | (A) Gene coverage across MCF-7, HepG2, HepaRG and undifferentiated iPSC cell lines **(B)** Over-represented processes exclusive for HepaRG and iPSC cells, analysed using WebGestalt.

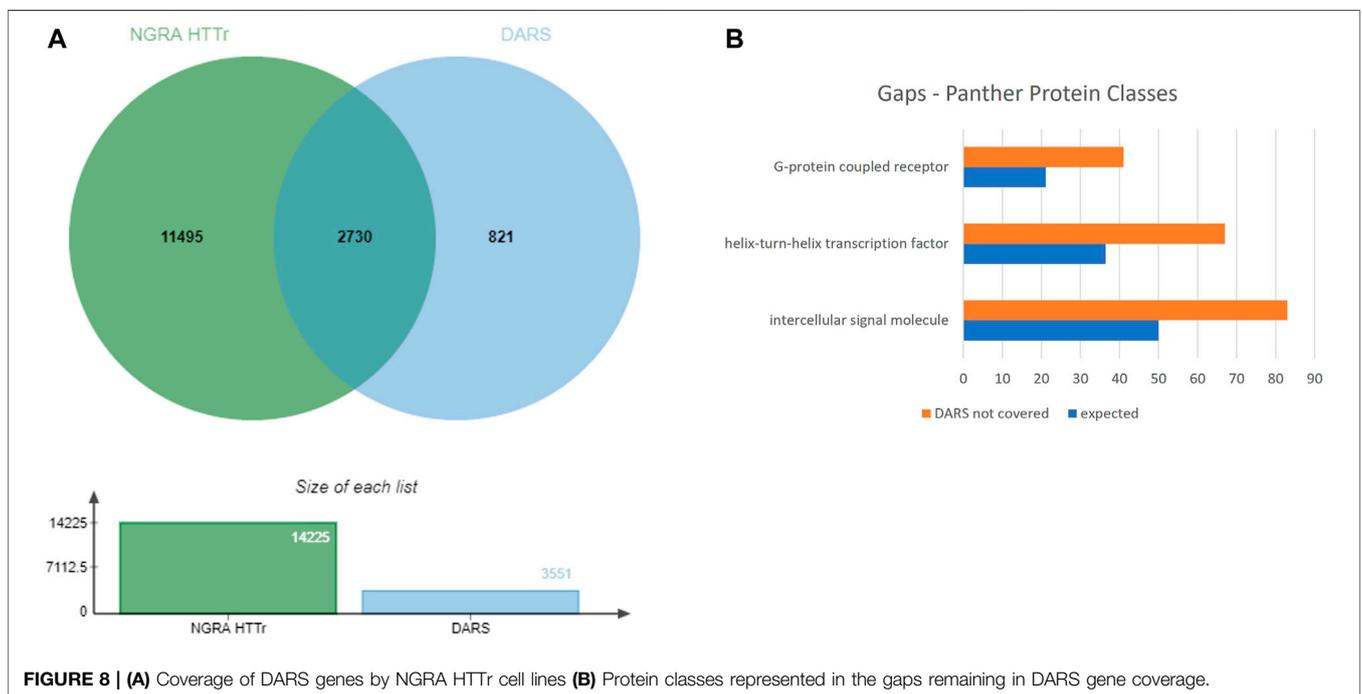


FIGURE 8 | (A) Coverage of DARS genes by NGRA HTTr cell lines **(B)** Protein classes represented in the gaps remaining in DARS gene coverage.

Some of the GO Biological Processes overrepresented in the genes unique to iPSC include neuropeptide signalling pathway, regulation of membrane potential, axon development and muscle system process (Figure 7B). Cytolysis, acute inflammatory response, and response to xenobiotic stimulus are some of the processes overrepresented in the genes unique

to the HepaRG dataset (Figure 7B). The number of unique genes in MCF7 and HepG2 was too small to reliably determine overrepresented processes in these cell lines.

The 3,551 DARS gene markers determined from the literature search were compared to the total of 14,225 genes found expressed in the MCF7, HepG2, HepaRG and iPSC cells. 2,730

TABLE 2 | DARS identified molecular process categorised depending on the cellular function.

Category	Examples
General cellular process	Signalling, DNA methylation, Cell differentiation, Homologous recombination, Cellular metabolic process, etc.
Specific cellular process	Retinol metabolic process, Myelination, Embryonic cleavage, Cytokine secretion, Meiotic cell cycle, etc.
General functional process	Cell migration, Bicellular tight junction assembly, Cell motility, etc.
Specific functional process	Sperm motility, Neuron migration, Axon guidance, Synapse assembly, Macrophage migration, etc.
Specific differentiation	T cell differentiation, Neurogenesis, Hepatocyte differentiation, Erythrocyte differentiation, Cardiocyte differentiation, etc.
Receptor or enzyme activity	1-phosphatidylinositol-3-kinase activity, MAP kinase activity, Interleukin-2 receptor activity, Fibroblast growth factor-mediated receptor activity, Cell adhesion mediator activity, etc.
Signalling pathway	Notch signalling pathway, Nodal signalling pathway, Hippo signalling, Protein kinase B signalling, Wnt signalling pathway, etc.
Cellular stress	Oxidative stress, Heat-shock response, Programmed cell death, Mitochondrial damage, Apoptotic process, etc.
Genotoxicity	Cell cycle checkpoint, Cellular response to DNA damage stimulus, DNA damage checkpoint, DNA repair, Mitotic DNA replication checkpoint, DNA integrity checkpoint, etc.

out of the 3,551 genes were found present in the gene set from the four cell lines (**Figure 8A**). In particular, transcriptional regulators, protein modifying enzymes, metabolite conversion enzymes, kinase and phosphatase modulators are overrepresented in the intersecting gene list. For the full list of overlapping and unique genes in the two datasets, refer to **Supplementary File S8**.

5.2.2 Biological Processes

The 474 DARS identified biological processes could be broadly classed into different categories depending on their role in cellular process (**Table 2**). Direct overlay of these biological processes to pathways or gene expression is difficult due to the fact that multiple signalling cascades may regulate these processes in a redundant or tissue-specific manner.

Several of the general cellular and functional processes would be part of housekeeping or homeostatic mechanisms within a cell and therefore, changes to some of these processes might be indicated through cell survival or cytotoxicity read outs.

The IPP panel, which assesses whether the chemical can interact with several GPCRs, enzymes, ion-channels, transporters, nuclear receptors, and bHLH transcription factors, covers about 13% of the receptor or enzyme activity related biological process key terms. This includes acetylcholine-gated channel activity, NMDA glutamate receptor activity, sodium or potassium channel related activity, serotonin receptor activity, and α_1/α_2 and β_1/β_2 adrenergic receptor activities.

The signalling pathways' key markers extracted as part of the DrBP biological processes largely overlap with the DARS genes and for all practical purposes are evaluated together for coverage and gaps.

Formation of the three germ layers and their derivatives are key developmental milestones and is the precursor to specific differentiation to various distinct cell fates (**Table 1**). This can be mimicked by the differentiation of pluripotent stem cells *in vitro*. The ReproTracker[®] assay uses human iPSCs differentiated into hepatocytes, cardiomyocytes or neural rosettes, which are representative of cell fates derived from each of the three germ layers; endoderm, mesoderm and ectoderm, measuring the expression of germ line (FOXA2, BMP4 and SOX1) and cell lineage (AFP, MYH6 and PAX6) specific markers genes in a time

dependent manner (see Methods). These marker genes are part of the identified DARS markers. However, the differentiation itself covers all signalling networks involved including transcriptomic changes (mRNA expression), posttranslational and epigenetic changes that play a role in the formation of hepatocytes, cardiomyocytes and neural rosettes (Gu et al., 2014; Vasconcellos et al., 2016; Ouyang and Wei, 2021). Assuming that these assays cover most of these three specified differentiation processes, transcriptomic analysis of the ReproTracker[®] derived tissues would allow a detailed analysis to define the overlap to the whole DARS gene set to get better understanding of the biological coverage of the assay.

The cell stress panel covered key cell stress markers, either as part of normal physiology ensuring survival or development or elimination of damaged or unwanted cells; or activated in response to xenobiotic exposure and stress.

The o/c amino acid ratio, as measured by the devTOXquickPredict[™] assay, is an indicator of metabolic health of cells and is shown to be a predictive signature of developmental toxicity (Palmer et al., 2013). Ornithine is a non-proteogenic amino acid that plays a role in several biochemical pathways and a decrease in cellular release reflects a perturbation in these pathways. Cystine is used by cells in glutathione production and a decreased uptake from the media is a likely indicator of change in cellular glutathione synthesis and redox balance. Thus, the devTOXquickPredict[™] assay covered key metabolic signatures of teratogenicity and based on this o/c ratio predicts the potential of a chemical to result in developmental toxicity (Palmer et al., 2013). Additionally, biochemical profiles reflecting MIEs or signalling pathways have been correlated with both its positive and negative predictions (Zurlinden et al., 2020). Zurlinden et al. used a logistic regression strategy considering ToxCast assay specific AC₅₀s together with a binary classification outcome model of the o/c ratio to identify sensitive and insensitive pathways associated with the assay predictions (Zurlinden et al., 2020). The annotated records for these MIEs flow largely into either RTK or GPCR signalling. These signalling pathways can be assessed to some extent using HTTr or iPP assays in the DART framework and can further provide mechanistic implications of the o/c ratio. This information can be explored or validated using the existing or additional NAMs.

Based on the training datasets (either publicly available or donated for training purposes) employed, some of the *in silico* tools issue a general alert for developmental toxicity or teratogenicity where as others indicate a specific mode of action, such as those related to the estrogen or androgen receptor signalling mechanisms. DEREK NEXUS includes alerts related to developmental toxicity, teratogenicity and testicular toxicity. The OECD QSAR Toolbox and VEGA both include alerts for developmental toxicants and estrogen receptor (ER) mediated effects. VEGA also includes alerts related to androgen receptor (AR). The OPERA Suite specifically covers estrogen and androgen receptor binding, agonism and antagonism. METEOR covers the prediction of metabolites for compounds in general, which can further be run through each of the above tools for any alert related to teratogenicity or ER and AR modules. MIE Atlas covers a range of alerts, including those related to ER and AR activities. The DART and mode of action related alerts arising from the *in silico* predictions provide a relevant direction of further testing and data generation using either iPP or ReproTracker[®] or the devTOXquickPredict[™] assays.

The biological processes that are related to genotoxicity are out of scope of the DART framework as potential for carcinogenicity would be covered as part of an overall safety assessment for a consumer product.

5.3 Gaps in Coverage of DARS Markers

Eighty-two out of the 3,551 genes were not expressed above the used threshold on expression levels in the four cell lines (see *Methods*). When the missing genes were classified into protein classes using Panther classification system, the over-represented classes were identified as GPCRs, helix-turn-helix (HTH) transcription factors and intercellular signalling molecules (**Figure 8B, Supplementary File S9**) (Mi et al., 2019). Of the 41 genes in the GPCR class, that lack coverage in the four cell lines, six of them can be evaluated through the iPP panels. Of the 67 HTH transcription factors, 61 are homeodomain transcription factors. The 83 intercellular signalling molecules can be further classified into chemokines, cytokines, growth factors, intercellular signalling molecules, interleukin superfamily members, neuropeptides, neurotrophic factor, and peptide hormones.

The specific cellular processes such as cytokine secretion or myelination or androgen biosynthesis cannot be evaluated entirely, given the current set of NAMs. Some of the specific cellular processes such as embryonic cleavage, although specific to developmental biology, may be similarly affected to generic cell division and can be evaluated together with genotoxic effects of chemicals that impact the cell cycle. However, if there are specific spindle poisons or cell division disrupting chemicals to which embryonic cells are more sensitive, those may not be adequately identified using the DART NAMs. Specific functional processes such as sperm motility or axon guidance or lymphocyte migration are also not within the scope of the current NAMs. Although some of cell adhesion related genes can be studied using the HTTr assay, which might add some weight of evidence to cell migratory functions. Differentiation processes such as T cell differentiation or oocyte maturation or Sertoli cell differentiation are also not

covered by the ReproTracker[®] assay or any other assays in the DART framework and will have to be addressed for certain modes of action. Some of the interleukins and cytokines or growth factor receptors or intracellular kinases are not covered by the iPP or CSP assays.

Thus, these gaps that still remain will have to be considered in the context of the mode of action of specific compounds or in a tissue-specific context. While some information in this context could be established from transcriptomic assays (such as differential regulation of genes involved in a biosynthetic pathway or those encoding growth factors), a complete coverage would require either more of the mechanistic data to add to the weight of evidence or data generated in an analogous model system. Similarly, studying specific functional processes may require higher tier models closer to the organoid or micro-physiological conditions. A tiered approach would have to be undertaken to generate additional data.

6 APPLICATION OF THE NGRA DART FRAMEWORK TO THE CAFFEINE CASE STUDY

6.1 Exposure Estimation

For each exposure scenario, the respective Maternal and Fetal C_{max} was derived. Following the PBK modelling workflow, non-pregnant PBK models were first constructed and verified against available human clinical data. Intravenous infusion data was used for model development by optimising the V_{max} parameter for the CYP1A2 pathway (**Supplementary Figure S1A**). Then the model was extrapolated into an oral PBK and a dermal PBK model by accounting for their corresponding absorption processes, which were verified by evaluating the model performance against available human clinical studies from oral and dermal exposure routes, respectively, which were not used for model developments (**Supplementary Figures S1B,C**). In these studies, the model's predictive performance was evaluated by visual inspection of the time-course curves, as well as calculating C_{max} ratios for observations over simulations. The observed/simulated ratios for C_{max} of all three clinical trials were well within 2-fold limit (calculation not shown), indicating the model predicted values are in good agreement with the respective observed values, therefore both the oral and dermal non-pregnant models were considered reasonable and validated (Jones et al., 2015).

The pregnancy model was further developed based on the non-pregnancy model which accounts for the physiological changes during pregnancy and validated against the plasma kinetics of caffeine after administration of 150 mg single dose in pregnant women. The simulated plasma concentration–time profile captured the observed PK parameters in great agreement (**Supplementary Figure S1D**). It is worth noting that the changes in maternal CYP1A2 activity have been captured in the software, which is the main enzyme responsible for most of caffeine metabolism. The good match between observed and predicted plasma time profile also indicates that metabolism of caffeine is well predicted too.

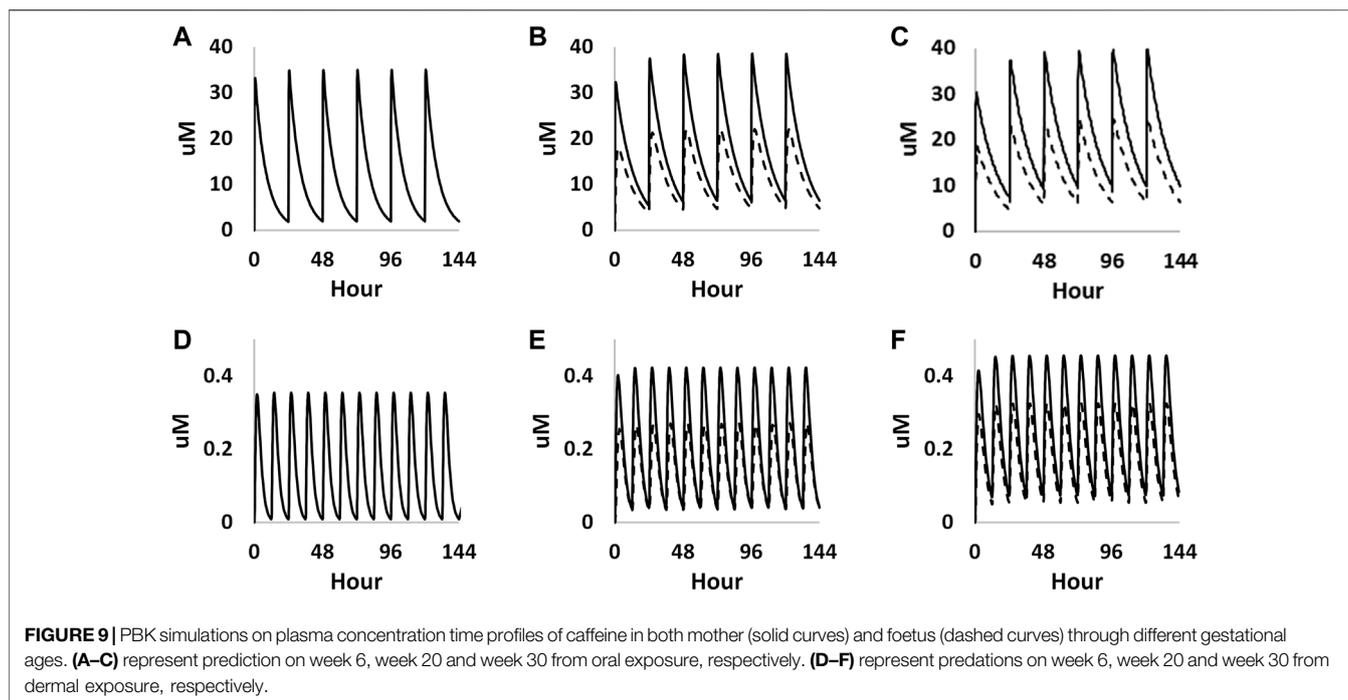


TABLE 3 | Summary of the predicted maternal and foetal plasma C_{max} of caffeine at steady state through different gestational ages from both oral and dermal exposure routes.

	Oral: 200 mg/day			Dermal: 0.1% caffeine in body lotion		
	Week 6	Week 20	Week 30	Week 6	Week 20	Week 30
Maternal plasma C_{max} (μM)	34.97	38.51	39.72	0.42	0.42	0.46
Foetal plasma C_{max} (μM)		22.02	25.27		0.27	0.32

The validated pregnant PBK model was then extended to consider the placental and foetal compartments during pregnancy to make predictions on both maternal and foetal concentrations from the two given exposure scenarios. As expected, caffeine exposure was significantly increased as a function of dose across the gestational ages, which is consistent with previous studies reporting a significant prolongation of caffeine elimination half-life in pregnant women.

The foetal exposure predictions were not verified due to lack of data on observed foetal concentrations in human clinical studies. The overall ratio between the predicted foetal and maternal (F/M) plasma C_{max} was in the range of 0.57–0.69 for the gestational age greater than 6 weeks (Figure 9 and Table 3). This F/M ratio aligns to F/M ratio after 60 min of perfusion of a term *ex vivo* human placenta (0.79) as well as the ratio reported in the EFSA report (Poulsen et al., 2009; Gundert-Remy 2015).

6.2 Collation of Existing Information

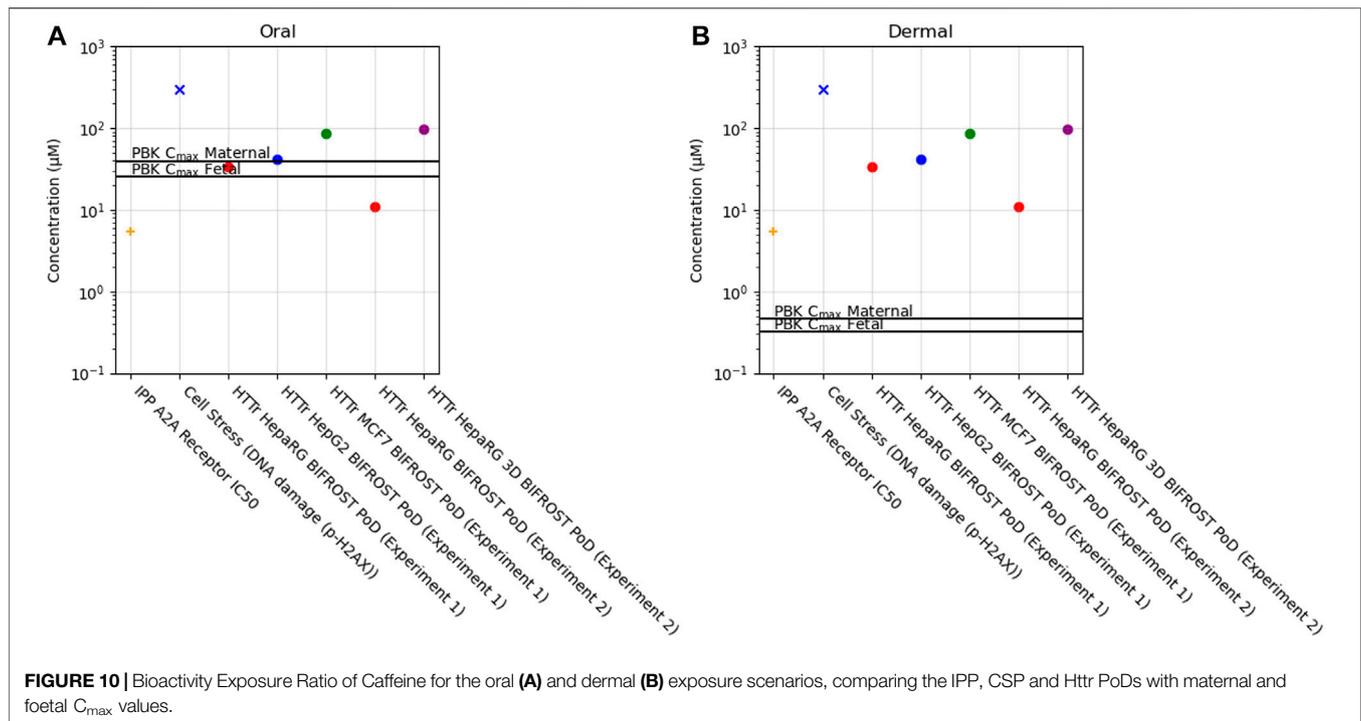
In silico tools predicted a wide range of possible toxicity effects based on caffeine's structure (Supplementary Table S4). DEREK, OECD toolbox and Vega predicted that caffeine has the potential to cause reproductive and developmental toxicity, however it should be noted that these models have been trained on caffeine's

animal experimental data. Other alerts included the potential for protein binding (MIE for skin sensitisation) and genotoxicity. These alerts were outside of the scope of this paper but in a real-life scenario these endpoints would be investigated. There were three positive alerts with high confidence predicted by the MIE ATLAS, namely the AChE (Acetylcholinesterase), ADORA2A (Adenosine receptor A2a) and PDE4A (Phosphodiesterase 4A). Notably, none of tools predicted binding to oestrogen or androgen receptors.

In order to ensure protection of human health, caffeine was experimentally investigated to determine whether it affects these endpoints at consumer relevant exposures using human-relevant tools (IPP, ReproTracker[®] and devTOXquickPredict[™]). In addition, CSP and HTTr were performed to increase the biological coverage given that the absence of *in silico* alerts for a given toxicity does not mean absence of effect.

6.3 *In Vitro* Biological Activity Characterisation

Of the 36 biomarkers in the cell stress panel, caffeine exposure only caused a borderline increase in the biomarker γ -H2AX (phosphorylated H2A histone family member X) (PoD =



304 μM) when tested up to 1,000 μM . Across the different cell models, PoDs derived from HTTr varied between 11 and 96 μM (Figure 10).

In the IPP, the only target hit was the adenosine 2A receptor out of the 60 targets tested (see *Methods* and **Supplementary File S4**). The follow-up dose-response analysis showed that the IC_{50} (C.I. 95%) for the A2A receptor was 5.3 μM (**Supplementary Figure S2**). Caffeine is a known non-selective antagonist of adenosine and specifically through the adenosine A2A receptor which is responsible for its psychoactive effect (Jacobson et al., 2020). The other targets predicted by MIE ATLAS were not identified by our panel when caffeine was tested at 10 μM . However, caffeine has also been reported to inhibit multiple phosphodiesterases (PDE1b, PDE2, PDE3, PDE4, PDE5), GABA_A (γ -aminobutyric acid receptor type A), and monoamine oxidase A (MAO-A) at concentrations greater than 10 μM and at not consumer relevant concentrations (Petzer et al., 2013; Pohanka, 2015; Jacobson et al., 2020).

Antagonism of adenosine receptors is a potential cause of concern for cardiac and neuronal tissue embryonic development among other adverse effects in adults (Rivkees et al., 2001; Rivkees and Wendler, 2012; Teerlink et al., 2012; Silva et al., 2013). The ReproTracker[®] assay included iPSCs that were differentiated towards cardiomyocytes, hepatocytes, and neuronal rosettes in the presence of caffeine. No signs of teratogenicity were observed up to concentrations of 100 μM . Similarly, caffeine did not affect the ornithine/cysteine ratio in the devTOXquickPredict[™] assay up to 500 μM .

In summary, the bioactivity data has shown that the most sensitive target was the adenosine 2A receptor.

6.4 Determination of Bioactivity Exposure Ratio (BER)

In this case study, we found that a daily consumption of 200 mg caffeine by pregnant women would lead to internal exposure (C_{max}) for both the mother and the foetus which was 10–20 fold greater than the most sensitive *in vitro* point of departure (adenosine 2A receptor activity) (Figure 10, **Supplementary Table S5**). Conversely, for the dermal exposure, the plasma C_{max} was around 10–20 fold lower than the PoD, suggesting that this exposure would not be sufficient to cause *in vivo* bioactivity. For the other biomarkers, BERs were higher but similar trend was observed between exposures, for the oral exposure BER ranged from 0.1 to 12 and for the dermal exposure between 12 and 950 across maternal and foetal predictions (**Supplementary Table S5**).

Previous studies have shown that a PoD based on the most sensitive pathway or biological response derived from NAMs provides a conservative estimate of the PoD *in vivo* (Thomas et al., 2013; Wetmore et al., 2013; Paul Friedman et al., 2020). In addition, the BER based on the adenosine 2A receptor antagonism would indicate that there is potential DART concern associated with the oral 200 mg/day exposure (BER < 1). However, because bioactivity may not result in an adverse response, the critical safety assessment question is whether this level of activity presents a risk to either the mother or the developing embryo or foetus. Follow-up *in vitro* studies in differentiating cardiomyocytes and neuronal rosettes, which are expected to express adenosine receptors (Headrick et al., 2013), were negative at caffeine exposure up to 100 μM , a concentration that is above the expected level in consumers

(~40–80 μM) and provide reassurance that these levels of activity may not be sufficient to adversely affect cardiac or neuronal development, at least in the periods covered by the assays. This conclusion is supported by the available epidemiology data, which suggest that the most sensitive adverse effect associated with oral caffeine exposure in pregnant women is intra-uterine growth retardation, and not specific malformations of the heart and brain (Gleason et al., 2021).

7 DISCUSSION

In this work, we have used the mechanistic knowledge of reproductive and developmental biology to evaluate the coverage and therefore, how fit-for-purpose the NAMs within our safety assessment framework are for DART endpoints. The “*in litero*” approach used to generate the master list of DARS markers combined the systematic categorization of reproduction and development into key stages and morphogenetic events along with use of specific query terms to tap into the existing knowledge base for each of the categories. A pooled master list of 3,551 genes, 474 cellular processes and 338 miRNAs was generated. Just the comparison of these DARS biomarkers genes with the read-outs of the baseline expression of genes in the four cell lines (HepG2, HepaRG, MCF7 and iPSC) and the IPP endpoints used in NAMs that are intended to be part of the DART framework indicated almost 80% coverage of DARS genes.

Of the remaining 821 DARS genes not covered in the basic expression profiles of the four cell lines from NAMs, the predominant classes of proteins are GPCRs, transcription factors with homeodomain transcription factors are the major class, chemokines, cytokines, growth factors, intercellular signalling molecules, interleukin superfamily members neuropeptides, neurotrophic factor, and peptide hormones. Although six of the GPCRs are covered by the iPP panel, the rest are presently not within the direct coverage of the DART framework assays.

Homeobox transcription factors (Hox) play a key role in antero-posterior and proximo-distal patterning in bilaterian embryos, and their expression is tightly regulated in development (Alexander et al., 2009; Montavon and Soshnikova, 2014). Their expression is closely linked to extension of antero-posterior body axis and controlled by signalling pathways such as Wnt, Fgf, and RA, as well as the Cdx transcription factors (Dubrulle et al., 2001; van den Akker et al., 2002; Aulehla et al., 2003; Moreno and Kintner, 2004). Similarly, BMP signalling has been shown to pattern the spinal neural tube by regulating Hox transcription factors and the roles of BMP, Shh, Fgf, TSH, IGF1 and TGF β are well recognized in regulating the thyroid transcription factors, including the Hox TFs, Nkx2-1 and Hhex (Timmer et al., 2002; Lopez-Marquez et al., 2021). Since the Hox genes are tightly and combinatorially regulated by growth factor signalling in a tissue specific manner, these upstream activating pathways may provide a better indication of regulation or perturbation of regulation in embryonic development. Indeed, the work by Franzosa et al. has shown that using Bayesian analysis, inferences

on MIEs and AOPs can be drawn from differential expression of transcription factors and can be used in an IATA-like approach (Franzosa et al., 2021). Further, transcriptomic analysis of the differentiated cell types from iPSCs as standardized into the ReproTracker[®] assay may provide germ layer specific or cell type specific niche landscape to probe for expression of these DARS biomarkers.

To assess an impact of chemical exposure on immune system related genes including cytokines, chemokines and interleukins, assays pertaining to its individual components can be considered. Lymphocytic and myeloid cells can be isolated from peripheral blood mononuclear cells (PBMCs) and can be tested to determine toxicity to each of these lineages specifically, upon chemical exposures (Hassan et al., 2007; Pessina and Bonomi, 2007). Other targeted functional assays for specific mechanisms include human lymphocyte activation (HuLA) assay, an antigen recall assay, similar to the *in vivo* T-cell-dependent antibody response (TDAR) where multiple immune cell types are needed to produce responses; multiple cytokines (IL-2, IFN- γ , IL-1 β , and IL-8) assay; and the BioMap panel of assays where test systems are constructed with one or more primary cell types from normal human donors stimulated with cytokines or growth factors recapitulate relevant signalling networks that naturally occur in human tissue or disease states and can be explored for specific biomarker readouts (Collinge et al., 2010; Kimura et al., 2018; Singer et al., 2019; Collinge et al., 2020).

Several of the DARS related biological processes are general enough to be indirectly covered either as a cell survival or cytoprotective mechanism or as steady state function to maintain stability or growth. Homologous recombination or endocytosis or DNA methylation would be such an inherent feature of all cells to a large extent. The same rationale can be extended to some of the housekeeping miRNA functionalities as well. However, certain tissue specific mechanisms cannot be completely delineated from the existing NAM models without further validation. A study conducted on human embryonic tissues outlined a high-resolution epigenomic atlas of human craniofacial development (Wilderman et al., 2018). Mapping such functional genomic profiles of embryonic tissues to transcriptomic profiles of cell lines, including iPSCs would further boost the confidence in the mechanistic coverage of NAMs (Haniiffa et al., 2021). Multi-generational processes are also challenging to study using generic cell lines with single cell types or those lacking meiotic features. For instance, studying parental exposure to compounds that may have an effect on embryo-foetal development in subsequent generation, such as sperm mediated developmental toxicity by means of epigenetic programming, including sperm DNA and histone modifications and non-coding RNAs in spermatozoa would require additional cell models with adequate mechanistic similarity or coverage (Bonde et al., 2016; Marcho et al., 2020). Furthermore, developing some of the tissue specific models have more inherent challenges than others and need additional considerations. For example, using the ReproTracker[®] set up to differentiate into additional cell fates such as osteoblast or other stromal lineages would be relatively straight forward compared to building an *in vitro* set up recapitulating the human testis (Oliver and Stukenborg, 2020).

Placental toxicity is another key cell type specific aspect of DART (Gingrich et al., 2020). However, the *ex vivo* or other advanced models of human placenta are still in their early stages of development (Hougaard, 2021).

Spatio-temporal and tissue-specific regulation is another critical aspect of developmental systems. This is key to the action of long-range signals such as growth factors and hormones and is well exemplified in regulation of both morphological (e.g., somitogenesis) and molecular processes (e.g., epigenetic modifications) (Maroto et al., 2012; He et al., 2020). While this work assesses the coverage of the key signalling and cellular machinery in reproduction and development, and the NAMs discussed do provide an opportunity to test time-dependent effects to certain extent, the entirety of spatio-temporal control as in a human embryo may not exist within a single NAM assay. The conservative outputs or PoDs from the NAMs described can be used directly in assessing risk, as a protective approach. However, these outputs cannot be directly related to the resulting cellular or morphological changes that may take place in a human embryo. Therefore, if a safety decision cannot be reached using this protective approach further targeted investigations or refinements will be needed. These may include the use of higher-tier approaches, such as hiPSC derived embryos combined with single cell omics techniques (Shahbazi et al., 2019; Zheng et al., 2019; Xiang et al., 2020; Sozen et al., 2021), bespoke microphysiological systems (MPS) and conceptual or agent-based or dynamic models (Kleinstreuer et al., 2013; Villeneuve et al., 2014; Capone et al., 2018; Thomas et al., 2019; Richards et al., 2020; Kim et al., 2021; Wiedenmann et al., 2021), to make a more explicit link between the cellular changes observed and an adverse outcome (a quantitative AOP). Such a bespoke approach will require a significant investment that may not be practical for day-to-day decision making.

MiRNAs have emerged over the last few years as important modulators of signalling networks in foetal development by regulating developmental transitions, lineage specification, and securing cellular identities (Ebert and Sharp, 2012; Ivey and Srivastava, 2015; Prodromidou and Matsas, 2019). Tissue specific microRNAs are found to be dynamically regulated in a cell lineage and developmental stage-specific manner (Vasconcellos et al., 2016; Ouyang and Wei, 2021). These dynamic changes can also be observed *in vitro*. Kim et al. showed specific temporal expression of miR-1, -30d, -133a, -143, -145, -378a, -499a during differentiation of human embryonic stem cells into a cardiac lineage (Kim et al., 2017), some of which were also identified in our literature search as important regulators for heart development (**Supplementary File S7**). So, while bioinformatic evaluation of the miRNA DARS markers was out of scope for this work we believe that miRNA expression is at least partially covered in all cell lines used as well as in the iPSC and the differentiating cells of the NAMs discussed for the DART framework. However, using miRNA analysis complementarily to mRNA analysis could bring further advantages for *in vitro* toxicity testing in future. Analogous to transcription factors, one miRNA can regulate several mRNAs influencing different functional processes at once. MIR-21 (second highest score in the literature extraction,

Supplementary File S7) is found to affect the expression levels of many different genes from different pathways and molecular events regulating epithelial-to-mesenchymal transition, a biological process essential for organ development (Huntley et al., 2018). While knockout of single miRNAs often produces subtle phenotypes under homeostatic conditions overexpression or misexpression of a single miRNA can promote remarkable alterations in cell differentiation and can even induce differentiation of multi- or pluripotent cells *in vitro* (Shenoy and Belloch, 2014; Channakkar et al., 2020; Jaafarpour et al., 2020). Dose-dependent perturbations of miRNA upon chemical exposures (Smirnova et al., 2014) and its link to the development of diseases for humans *in vivo* have been reported (Schraml et al., 2017; Wallace et al., 2020). Especially the fact that circulating miRNAs in the maternal blood can be used as biomarkers for pregnancy complications and for detecting foetal abnormalities (Yang et al., 2020) make them important clinical biomarkers and could also help identify new mode of actions for developmental toxicity.

A workshop reviewing the application of the nine ICCR principles to NGRA case studies and how can these principles build confidence in safety decision making, using NAMs, identified the evaluation of the biological coverage as one of the seven areas that would help to make NGRA useful for cosmetic ingredients (Dent et al., 2021). Either using a multitude of cell and tissue types or focussing on key pathways and mechanisms of concern irrespective of the origin of the cell type are two approaches that can be used alternatively or in combination to address safety concerns. The current approach of generating a reference list of DARS biomarkers enables such a robust evaluation of DART related biological coverage. Overall, this evaluation found the NAMs to broadly cover the key signalling processes and mechanisms underlying the apical endpoints in traditional DART studies, such as those specified within the OECD guidelines and undertaken for regulatory purposes, thereby enabling their assessment in a tiered manner to be protective of human health. This approach also provides the opportunity to identify gaps and address them through the above-mentioned propositions of either additional AOP-based or cell type specific analysis. Furthermore, the individual list of DARS markers available for each of the stages can be used as either foundational or supporting knowledge for DART endpoint-specific AOPs, thereby enabling a refinement in NGRA.

One of the objectives of this paper was to illustrate how the NAMs within the Developmental and Reproductive Safety Framework can be protective of DART effects spanning the key stages in reproduction and development. We followed this framework for two hypothetical exposure scenarios, applying the PBK model approaches to estimate caffeine's internal maternal and foetal exposure, combined with the biological activity characterisation using the NAMS evaluated in the biological coverage section. Caffeine effects during pregnancy have been a topic of debate for many years, but safety agencies around the world agree that a daily consumption of 200 mg is unlikely to lead to significant adverse reproductive and developmental effects (USDA 2015; EFSA, 2015; FDA, 2018; ACOG, 2020; Gleason

et al., 2021; James, 2021; NHS, 2020). Here we found that an oral dose below 20 mg/day and a dermal exposure with 0.1% caffeine would result in a maternal BER higher than one implying that this exposure would result in no *in vivo* biological activity. It should be noted that this exposure does not take into account any uncertainties associated with e.g. inter-individual variability in the pharmacodynamics and pharmacokinetics (Nehlig, 2018). However, given the existing knowledge of caffeine, this level would clearly be protective of human health. This case study focused on the risk assessment of caffeine without considering the potential exposure and toxicity of its metabolites. We have, however, incorporated the reduced caffeine metabolism observed during pregnancy (decrease in CYP1A2 activity) which leads to higher caffeine plasma concentrations (Tracy et al., 2005). As a further refinement, the risk assessment should consider the biologically active metabolites, paraxanthine, theophylline, and theobromine (Institute of Medicine et al., 2001; Orru et al., 2013).

The caffeine case study is a practical example that demonstrates how to integrate relevant NAMs for safety decisions without generating new animal data and in this instance, the tools provided a conservative estimate of risk. However, teratogens and reproductive toxicants are known to act through multiple mechanisms of toxicity (Mattison and Thomford, 1989; van Gelder et al., 2014) which means that our NAM toolbox might not provide adequate biological coverage to all chemicals. Therefore, to build confidence that NAMs can be protective of DART effects for a wide range of chemicals, we need to generate larger datasets for multiple types of chemistries, toxicity modes of action and exposure scenarios. Work is ongoing to generate the NAM toolbox data for a substantial number of known human teratogens and non-teratogens with known human exposure. This approach can be used not only to evaluate how protective the toolbox is but also to identify potential areas of refinement (e.g., application of higher tier tools triggered by molecular signatures or specific targets) and to develop new tools (e.g., to address gaps in the biological coverage). Formulating more and more DART specific case studies and using the NAMs data to guide the evaluation of novel chemicals will help realise the ambition of NGRA and non-animal alternatives for safety decision making.

Lastly, we are yet to have a unified database as a resource that can share the available information on epidemiological and environmental aspects or clinical biomarkers for the large number of industrial chemicals and pharmaceutical compounds (Hougaard, 2021). Such a database would be

immensely valuable in validating and benchmarking NAMs outputs with the gold standard human evidence. A clinical study of prenatal exome sequencing analysis, after excluding for extra chromosomes or large deletions or duplications of chromosomes, identified a list of genes associated with foetal structural anomalies (Lord et al., 2019). Roughly 50% of these genes overlapped with the DARS markers which was a good validation of the DARS list as not all alleles associated with prenatal defects resulted in a loss of function. Indeed, a collaboration between toxicologists, epidemiologists and clinicians studying birth defects and reproductive disorders would strengthen and expedite the validation and application of NAMs towards safety of human health. A workshop or a similar platform to enable such cross-talks would be very valuable in the future.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/**Supplementary Material**, further inquiries can be directed to the corresponding author.

AUTHOR CONTRIBUTIONS

RR and PK have contributed equally to the paper. All authors have made a substantial and intellectual contribution to the scientific work and drafting of the paper.

ACKNOWLEDGMENTS

The authors would like to thank Bobbie Bradford, Alistair Middleton, Carl Westmoreland, Stella Cochrane, Dave Allen, Agnus Karmaus, and Shannon Bell for their helpful technical advice and comments during the preparation of this manuscript.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/ftox.2022.838466/full#supplementary-material>

REFERENCES

- ACOG (2020). Moderate Caffeine Consumption during Pregnancy. Available at: <https://www.acog.org/-/media/project/acog/acogorg/clinical/files/committee-opinion/articles/2010/08/moderate-caffeine-consumption-during-pregnancy.pdf> (Accessed November 10 2021).
- Alexander, T., Nolte, C., and Krumlauf, R. (2009). Hox Genes and Segmentation of the Hindbrain and Axial Skeleton. *Annu. Rev. Cel Dev. Biol.* 25, 431–456. doi:10.1146/annurev.cellbio.042308.113423
- Alles, J., Fehlmann, T., Fischer, U., Backes, C., Galata, V., Minet, M., et al. (2019). An Estimate of the Total Number of True Human miRNAs. *Nucleic Acids Res.* 47 (7), 3353–3364. doi:10.1093/nar/gkz097
- Ankley, G. T., Bennett, R. S., Erickson, R. J., Hoff, D. J., Hornung, M. W., Johnson, R. D., et al. (2010). Adverse Outcome Pathways: a Conceptual Framework to Support Ecotoxicology Research and Risk Assessment. *Environ. Toxicol. Chem.* 29 (3), 730–741. doi:10.1002/etc.34
- Arnesdotter, E., Spinu, N., Firman, J., Ebbrell, D., Cronin, M. T. D., Vanhaecke, T., et al. (2021). Derivation, Characterisation and Analysis of an Adverse Outcome Pathway Network for Human Hepatotoxicity. *Toxicology* 459, 152856. doi:10.1016/j.tox.2021.152856
- Ashburner, M., Ball, C. A., Blake, J. A., Botstein, D., Butler, H., Cherry, J. M., et al. (2000). Gene Ontology: Tool for the Unification of Biology. *Nat. Genet.* 25 (1), 25–29. doi:10.1038/75556
- Aulehla, A., Wehrle, C., Brand-Saberi, B., Kemler, R., Gossler, A., Kanzler, B., et al. (2003). Wnt3a Plays a Major Role in the Segmentation Clock Controlling

- Somitogenesis. *Develop. Cel* 4 (3), 395–406. doi:10.1016/s1534-5807(03)00055-8
- Baker, N., Knudsen, T., and Williams, A. J. (2017). Abstract Sifter: a Comprehensive Front-End System to PubMed. *F1000Res* 6, 2164. doi:10.12688/f1000research.12865.1
- Baltazar, M. T., Cable, S., Carmichael, P. L., Cubberley, R., Cull, T., Delagrang, M., et al. (2020). A Next-Generation Risk Assessment Case Study for Coumarin in Cosmetic Products. *Toxicol. Sci.* 176 (1), 236–252. doi:10.1093/toxsci/kfaa048
- Berggren, E., White, A., Ouedraogo, G., Paini, A., Richarz, A.-N., Bois, F. Y., et al. (2017). Ab Initio chemical Safety Assessment: A Workflow Based on Exposure Considerations and Non-animal Methods. *Comput. Toxicol.* 4, 31–44. doi:10.1016/j.comtox.2017.10.001
- Blanchard, J., and Sawers, S. J. A. (1983). The Absolute Bioavailability of Caffeine in Man. *Eur. J. Clin. Pharmacol.* 24 (1), 93–98. doi:10.1007/BF00613933
- Bonde, J. P., Flachs, E. M., Rimborg, S., Glazer, C. H., Giwercman, A., Ramlau-Hansen, C. H., et al. (2016). The Epidemiologic Evidence Linking Prenatal and Postnatal Exposure to Endocrine Disrupting Chemicals with Male Reproductive Disorders: a Systematic Review and Meta-Analysis. *Hum. Reprod. Update* 23 (1), 104–125. doi:10.1093/humupd/dmw036
- Bowes, J., Brown, A. J., Hamon, J., Jarolimek, W., Sridhar, A., Waldron, G., et al. (2012). Reducing Safety-Related Drug Attrition: the Use of *In Vitro* Pharmacological Profiling. *Nat. Rev. Drug Discov.* 11 (12), 909–922. doi:10.1038/nrd3845
- Brazier, J. L., Ritter, J., Berland, M., Khenfer, D., and Faucon, G. (1983). Pharmacokinetics of Caffeine during and after Pregnancy. *Dev. Pharmacol. Ther.* 6 (5), 315–322. doi:10.1159/000457332
- Capone, C., Gigante, G., and Del Giudice, P. (2018). Spontaneous Activity Emerging from an Inferred Network Model Captures Complex Spatio-Temporal Dynamics of Spike Data. *Sci. Rep.* 8 (1), 17056. doi:10.1038/s41598-018-35433-0
- Channakkar, A. S., Singh, T., Pattnaik, B., Gupta, K., Seth, P., and Adlakha, Y. K. (2020). MiRNA-137-mediated Modulation of Mitochondrial Dynamics Regulates Human Neural Stem Cell Fate. *Stem Cells* 38 (5), 683–697. doi:10.1002/stem.3155
- Cheng, W. S. C., Murphy, T. L., Smith, M. T., Cooksley, W. G. E., Halliday, J. W., and Powell, L. W. (1990). Dose-dependent Pharmacokinetics of Caffeine in Humans: Relevance as a Test of Quantitative Liver Function. *Clin. Pharmacol. Ther.* 47 (4), 516–524. doi:10.1038/clpt.1990.66
- Codaccioni, M., Bois, F., and Brochot, C. (2019). Placental Transfer of Xenobiotics in Pregnancy Physiologically-Based Pharmacokinetic Models: Structure and Data. *Comput. Toxicol.* 12, 100111. doi:10.1016/j.comtox.2019.100111
- Collinge, M., Cole, S. H., Schneider, P. A., Donovan, C. B., Kamperschroer, C., and Kawabata, T. T. (2010). Human Lymphocyte Activation Assay: Anin Vitromethod for Predictive Immunotoxicity Testing. *J. Immunotoxicol.* 7 (4), 357–366. doi:10.3109/1547691X.2010.523881
- Collinge, M., Schneider, P., Li, D., Parish, S., Dumont, C., Freebern, W., et al. (2020). Cross-company Evaluation of the Human Lymphocyte Activation Assay. *J. Immunotoxicol.* 17 (1), 51–58. doi:10.1080/1547691X.2020.1725694
- Dent, M., Amaral, R. T., Da Silva, P. A., Ansell, J., Boisleve, F., Hatao, M., et al. (2018). Principles Underpinning the Use of New Methodologies in the Risk Assessment of Cosmetic Ingredients. *Comput. Toxicol.* 7, 20–26. doi:10.1016/j.comtox.2018.06.001
- Dent, M. P., Vaillancourt, E., Thomas, R. S., Carmichael, P. L., Ouedraogo, G., Kojima, H., et al. (2021). Paving the Way for Application of Next Generation Risk Assessment to Safety Decision-Making for Cosmetic Ingredients. *Regul. Toxicol. Pharmacol.* 125, 105026. doi:10.1016/j.yrtph.2021.105026
- Dhyani, A. (2021). Placenta Geometry in the 2nd Trimester Assessed with Magnetic Resonance Imaging. Available at: <https://justjournal.org/2021/04/26/placenta-geometry-in-the-2nd-trimester-assessed/> (Accessed October 22 2021).
- Dubrulle, J., McGrew, M. J., and Pourquié, O. (2001). FGF Signaling Controls Somite Boundary Position and Regulates Segmentation Clock Control of Spatiotemporal Hox Gene Activation. *Cell* 106 (2), 219–232. doi:10.1016/s0092-8674(01)00437-8
- Ebert, M. S., and Sharp, P. A. (2012). Roles for microRNAs in Conferring Robustness to Biological Processes. *Cell* 149 (3), 515–524. doi:10.1016/j.cell.2012.04.005
- ECC/HC (2016). *Chemicals Management Plan (CMP) Science Committee Objectives Paper Meeting No. 5—Integrating New Approach Methodologies within the CMP: Identifying Priorities for Risk Assessment, Existing Substances Risk Assessment Program*. Ottawa, Ontario, Canada: Government of Canada. Available at: <http://www.ec.gc.ca/ese-ees/default.asp?lang=En&n=172614CE-1>.
- ECHA (2016). “New Approach Methodologies in Regulatory Science,” in Proceedings of a Scientific Workshop, Helsinki, Finland, 19–20 April, 2016. Available at: https://echa.europa.eu/documents/10162/22816069/scientific_ws_proceedings_en.pdf.
- ECHA (2017). *ECHA Strategic Plan 2019–2023*. Helsinki, Finland: ECHA. Available at: https://echa.europa.eu/documents/10162/26075800/echa_strategic_plan_2019-2023_en.pdf/3457ccff-7240-2c1f-3a15-fa6e5e65ac56.
- EFSA (2015). Scientific Opinion on the Safety of Caffeine. Available at: <https://efsa.onlinelibrary.wiley.com/doi/pdf/10.2903/j.efsa.2015.4102> (Accessed November 10 2021).
- EPA (2018). *Final Strategic Plan to Promote Development and Implementation of Alternative Test Methods Supporting Toxic Substances Control Act*, 83. Washington, DC: Federal Register, 30167–30168.
- FDA (2018). Available at: <https://www.fda.gov/consumers/consumer-updates/spilling-beans-how-much-caffeine-too-much>.
- Franzosa, J. A., Bonzo, J. A., Jack, J., Baker, N. C., Kothiya, P., Witek, R. P., et al. (2021). High-throughput Toxicogenomic Screening of Chemicals in the Environment Using Metabolically Competent Hepatic Cell Cultures. *NPJ Syst. Biol. Appl.* 7 (1), 7. doi:10.1038/s41540-020-00166-2
- Gabbe, S., Niebyl, J., Simpson, J., Landon, M., Galan, H., Jauniaux, E., et al. (2017). “Chapter 2 - Fetal Development and Physiology,” in *Obstetrics: Normal and Problem Pregnancies*. 7th ed (Amsterdam, Netherlands: Elsevier).
- Gilbert, S. F., and Barresi, M. J. F. (2016). *Developmental Biology*. 11 ed. Sunderland, MA: Sinauer Associates.
- Gingrich, J., Ticiani, E., and Veiga-Lopez, A. (2020). Placenta Disrupted: Endocrine Disrupting Chemicals and Pregnancy. *Trends Endocrinol. Metab.* 31 (7), 508–524. doi:10.1016/j.tem.2020.03.003
- Gleason, J. L., Tekola-Ayele, F., Sundaram, R., Hinkle, S. N., Vafai, Y., Buck Louis, G. M., et al. (2021). Association between Maternal Caffeine Consumption and Metabolism and Neonatal Anthropometry. *JAMA Netw. Open* 4 (3), e213238. doi:10.1001/jamanetworkopen.2021.3238
- Griffiths, S. K., and Campbell, J. P. (2015). Placental Structure, Function and Drug transfer Continuing Education in Anaesthesia Critical Care & Pain. *Continuing Educ. Anaesth. Crit. Care Pain* 15 (2), 84–89. doi:10.1093/bjaceaccp/mku013
- Grzegorzewski, J., Bartsch, F., Köller, A., and König, M. (2021). Pharmacokinetics of Caffeine: A Systematic Analysis of Reported Data for Application in Metabolic Phenotyping and Liver Function Testing. *bioRxiv*, 452094. doi:10.1101/2021.07.12.452094
- Gu, Y., Liu, G.-H., Plongthongkum, N., Benner, C., Yi, F., Qu, J., et al. (2014). Global DNA Methylation and Transcriptional Analyses of Human ESC-Derived Cardiomyocytes. *Protein Cell* 5 (1), 59–68. doi:10.1007/s13238-013-0016-x
- Gundert-Remy, U. (2015). “Pharmacokinetic and Pharmacodynamic Effects of Caffeine,” in Efsa Stakeholders Meeting on the Safety Of Caffeine Brussels, Brussels, Belgium, March 5, 2015. Available at: <https://www.efsa.europa.eu/en/events/event/caffeine-efsa-meets-stakeholders>.
- Haniffa, M., Taylor, D., Linnarsson, S., Aronow, B. J., Bader, G. D., Barker, R. A., et al. (2021). Human Cell Atlas Developmental Biological, NA Roadmap for the Human Developmental Cell Atlas. *Nature* 597 (7875), 196–205. doi:10.1038/s41586-021-03620-1
- Hassan, S. B., Haglund, C., Aˆleskog, A., Larsson, R., and Lindhagen, E. (2007). Primary Lymphocytes as Predictors for Species Differences in Cytotoxic Drug Sensitivity. *Toxicol. Vitro* 21 (6), 1174–1181. doi:10.1016/j.tiv.2007.03.009
- Hatherell, S., Baltazar, M. T., Reynolds, J., Carmichael, P. L., Dent, M., Li, H., et al. (2020). Identifying and Characterizing Stress Pathways of Concern for Consumer Safety in Next-Generation Risk Assessment. *Toxicol. Sci.* 176 (1), 11–33. doi:10.1093/toxsci/kfaa054
- He, J., Zheng, Z., Luo, X., Hong, Y., Su, W., and Cai, C. (2020). Histone Demethylase PHF8 Is Required for the Development of the Zebrafish Inner Ear and Posterior Lateral Line. *Front. Cel Dev. Biol.* 8, 566504. doi:10.3389/fcell.2020.566504

- Headrick, J. P., Ashton, K. J., Rosemeyer, R. B., and Peart, J. N. (2013). Cardiovascular Adenosine Receptors: Expression, Actions and Interactions. *Pharmacol. Ther.* 140 (1), 92–111. doi:10.1016/j.pharmthera.2013.06.002
- Hougaard, K. S. (2021). Next Generation Reproductive and Developmental Toxicology: Crosstalk into the Future. *Front. Toxicol.* 3, 652571. doi:10.3389/ftox.2021.652571
- Huntley, R. P., Kramarz, B., Sawford, T., Umrao, Z., Kalea, A., Acquah, V., et al. (2018). Expanding the Horizons of microRNA Bioinformatics. *RNA* 24 (8), 1005–1017. doi:10.1261/rna.065565.118
- ICCR (2018). Report on Integrated Strategies for Safety Assessment of Cosmetic Ingredients: Part 2. Available at: https://www.iccr-cosmetics.org/files/8315/4322/3079/ICCR_Integrated_Strategies_for_Safety_Assessment_of_Cosmetic_Ingredients_Part_2.pdf (Accessed July 30 2021).
- Institute of Medicine, Food and Nutrition Board, Committee on Military Nutrition Research (2001). *Caffeine for the Sustainment of Mental Task Performance: Formulations for Military Operations*. Washington (DC): National Academies Press, US.
- Ivey, K. N., and Srivastava, D. (2015). microRNAs as Developmental Regulators. *Cold Spring Harb Perspect. Biol.* 7 (7), a008144. doi:10.1101/cshperspect.a008144
- Jaafarpour, Z., Soleimani, M., Hosseinkhani, S., Geramizadeh, B., Yaghmaei, P., Mobarra, N., et al. (2020). Overexpression of microRNA-375 and microRNA-122 Promotes the Differentiation of Human Induced Pluripotent Stem Cells into Hepatocyte-like Cells. *Biologicals* 63, 24–32. doi:10.1016/j.biologicals.2019.12.005
- Jacobson, K. A., Gao, Z. G., Matricon, P., Eddy, M. T., and Carlsson, J. (2020). Adenosine A 2A Receptor Antagonists: from Caffeine to Selective Non-xanthines. *Br. J Pharmacol.* doi:10.1111/bph.15103
- James, J. E. (2021). Maternal Caffeine Consumption and Pregnancy Outcomes: a Narrative Review with Implications for Advice to Mothers and Mothers-To-Be. *BMJ EBM* 26 (3), 114–115. doi:10.1136/bmjebm-2020-111432
- Jassal, B., Matthews, L., Viteri, G., Gong, C., Lorente, P., Fabregat, A., et al. (2020). The Reactome Pathway Knowledgebase. *Nucleic Acids Res.* 48 (D1), D498–D503. doi:10.1093/nar/gkz1031
- Ji, R., Cheng, Y., Yue, J., Yang, J., Liu, X., Chen, H., et al. (2007). MicroRNA Expression Signature and Antisense-Mediated Depletion Reveal an Essential Role of MicroRNA in Vascular Neointimal Lesion Formation. *Circ. Res.* 100 (11), 1579–1588. doi:10.1161/CIRCRESAHA.106.141986
- Johnstone, R. H., Bardenet, R., Gavaghan, D. J., and Mirams, G. R. (2017). Hierarchical Bayesian Inference for Ion Channel Screening Dose-Response Data [version 2; Peer Review: 2 Approved]. Wellcome Open Research. Available at: <https://wellcomeopenresearch.org/articles/1-6/v2> (Accessed September 14 2021).
- Jones, H., Chen, Y., Gibson, C., Heimbach, T., Parrott, N., Peters, S., et al. (2015). Physiologically Based Pharmacokinetic Modeling in Drug Discovery and Development: A Pharmaceutical Industry Perspective. *Clin. Pharmacol. Ther.* 97 (3), 247–262. doi:10.1002/cpt.37
- Kim, Y. Y., Min, H., Kim, H., Choi, Y. M., Liu, H. C., and Ku, S.-Y. (2017). Differential MicroRNA Expression Profile of Human Embryonic Stem Cell-Derived Cardiac Lineage Cells. *Tissue Eng. Regen. Med.* 14 (2), 163–169. doi:10.1007/s13770-017-0051-4
- Kim, J., Lee, K.-T., Lee, J. S., Shin, J., Cui, B., Yang, K., et al. (2021). Fungal Brain Infection Modelled in a Human-Neurovascular-Unit-On-A-Chip with a Functional Blood-Brain Barrier. *Nat. Biomed. Eng.* 5 (8), 830–846. doi:10.1038/s41551-021-00743-8
- Kimura, Y., Fujimura, C., Ito, Y., Takahashi, T., Terui, H., and Aiba, S. (2018). Profiling the Immunotoxicity of Chemicals Based on *In Vitro* Evaluation by a Combination of the Multi-ImmunoTox Assay and the IL-8 Luc Assay. *Arch. Toxicol.* 92 (6), 2043–2054. doi:10.1007/s00204-018-2199-7
- Kleinstreuer, N., Dix, D., Rountree, M., Baker, N., Sipes, N., Reif, D., et al. (2013). A Computational Model Predicting Disruption of Blood Vessel Development. *Plos Comput. Biol.* 9 (4), e1002996. doi:10.1371/journal.pcbi.1002996
- Kliegman, R. M., Geme, J. W., Blum, N. J., Shah, S. S., Tasker, R. C., and Wilson, K. M. (2020). “Chapter 20 - Assessment of Fetal Growth and Development,” in *Nelson Textbook of Pediatrics*. 21 ed (Amsterdam, Netherlands: Elsevier).
- Knapen, D., Vergauwen, L., Villeneuve, D. L., and Ankley, G. T. (2015). The Potential of AOP Networks for Reproductive and Developmental Toxicity Assay Development. *Reprod. Toxicol.* 56, 52–55. doi:10.1016/j.reprotox.2015.04.003
- Kozomara, A., Birgaoanu, M., and Griffiths-Jones, S. (2019). miRBase: from microRNA Sequences to Function. *Nucleic Acids Res.* 47 (D1), D155–D162. doi:10.1093/nar/gky1141
- Krämer, A., Green, J., Pollard, J., Jr., and Tugendreich, S. (2014). Causal Analysis Approaches in Ingenuity Pathway Analysis. *Bioinformatics* 30 (4), 523–530. doi:10.1093/bioinformatics/btt703
- López-Márquez, A., Carrasco-López, C., Fernández-Méndez, C., and Santisteban, P. (2021). Unraveling the Complex Interplay between Transcription Factors and Signaling Molecules in Thyroid Differentiation and Function, from Embryos to Adults. *Front. Endocrinol.* 12, 654569. doi:10.3389/fendo.2021.654569
- Lord, J., McMullan, D. J., Eberhardt, R. Y., Rinck, G., Hamilton, S. J., Quinlan-Jones, E., et al. (2019). Prenatal Exome Sequencing Analysis in Fetal Structural Anomalies Detected by Ultrasonography (PAGE): a Cohort Study. *Lancet* 393 (10173), 747–757. doi:10.1016/S0140-6736(18)31940-8
- Mallon, B. S., Chenoweth, J. G., Johnson, K. R., Hamilton, R. S., Tesar, P. J., Yavatkar, A. S., et al. (2013). StemCellDB: the Human Pluripotent Stem Cell Database at the National Institutes of Health. *Stem Cell Res.* 10 (1), 57–66. doi:10.1016/j.scr.2012.09.002
- Marcho, C., Oluwayiose, O. A., and Pilsner, J. R. (2020). The Preconception Environment and Sperm Epigenetics. *Andrology* 8 (4), 924–942. doi:10.1111/andr.12753
- Maroto, M., Bone, R. A., and Dale, J. K. (2012). Somitogenesis. *Development* 139 (14), 2453–2456. doi:10.1242/dev.069310
- Martens, M., Ammar, A., Riutta, A., Waagmeester, A., Slenker, D. N., Hanspers, K., et al. (2020). WikiPathways: Connecting Communities. *Nucleic Acids Res.* 49, D613–D621. doi:10.1093/nar/gkaa1024
- Mattison, D. R., and Thomford, P. J. (1989). The Mechanisms of Action of Reproductive Toxicants. *Toxicol. Pathol.* 17 (2), 364–376. doi:10.1177/019262338901700213
- Meek, M. E., Boobis, A., Cote, I., Dellarco, V., Fotakis, G., Munn, S., et al. (2014). New Developments in the Evolution and Application of the WHO/IPCS Framework on Mode of Action/species Concordance Analysis. *J. Appl. Toxicol.* 34 (1), 1–18. doi:10.1002/jat.2949
- Mi, H., Muruganujan, A., Huang, X., Ebert, D., Mills, C., Guo, X., et al. (2019). Protocol Update for Large-Scale Genome and Gene Function Analysis with the PANTHER Classification System (v.14.0). *Nat. Protoc.* 14 (3), 703–721. doi:10.1038/s41596-019-0128-8
- Montavon, T., and Soshnikova, N. (2014). Hox Gene Regulation and Timing in Embryogenesis. *Semin. Cell Develop. Biol.* 34, 76–84. doi:10.1016/j.semcdb.2014.06.005
- Moreno, T. A., and Kintner, C. (2004). Regulation of Segmental Patterning by Retinoic Acid Signaling during *Xenopus* Somitogenesis. *Develop. Cel* 6 (2), 205–218. doi:10.1016/s1534-5807(04)00026-7
- National Research Council, Commission on Life Sciences, Board on Environmental Studies and Toxicology, Committee on Developmental Toxicology (2000). *Scientific Frontiers in Developmental Toxicology and Risk Assessment*. Washington (DC): National Academies Press.
- Nehlig, A. (2018). Interindividual Differences in Caffeine Metabolism and Factors Driving Caffeine Consumption. *Pharmacol. Rev.* 70 (2), 384–411. doi:10.1124/pr.117.014407
- NHS (2020). NHS Advice. Available at: <https://www.nhs.uk/pregnancy/keeping-well/foods-to-avoid/> (Accessed November 10 2021).
- OECD (2005). Guidance Document on the Validation and International Acceptance of New or Updated Test Methods for hazard Assessment: OECD Series on Testing and Assessment, Number 34 (ENV/JM/MONO(2005)14). Available at: [https://www.oecd.org/officialdocuments/publicdisplaydocumentpdf/?cote=ENV/JM/MONO\(2018\)19&doLanguage=en](https://www.oecd.org/officialdocuments/publicdisplaydocumentpdf/?cote=ENV/JM/MONO(2018)19&doLanguage=en) (Accessed July 30 2021).
- OECD (2017). Guidance Document for Describing Non-guideline *In Vitro* Test Methods, OECD Series on Testing and Assessment, No. 211 (ENV/JM/MONO(2014)35). Available at: <https://www.oecd-ilibrary.org/docserver/9789264274730-en.pdf?expires=1626436326&id=id&accname=guest&checksum=BF5B058705B87255A946A52AC4BE4984> (Accessed July 30 2021).
- OECD (2018). Guidance Document on Good *In Vitro* Method Practices (GIVIMP). OECD Series on Testing and Assessment, No. 286. doi:10.1787/9789264304796-en

- Oliver, E., and Stukenborg, J. B. (2020). Rebuilding the Human Testis *In Vitro*. *Andrology* 8 (4), 825–834. doi:10.1111/andr.12710
- Orrú, M., Guitart, X., Karcz-Kubicha, M., Solinas, M., Justinova, Z., Barodia, S. K., et al. (2013). Psychostimulant Pharmacological Profile of Paraxanthine, the Main Metabolite of Caffeine in Humans. *Neuropharmacology* 67, 476–484. doi:10.1016/j.neuropharm.2012.11.029
- Otberg, N., Patzelt, A., Rasulev, U., Hagemeister, T., Linscheid, M., Sinkgraven, R., et al. (2008). The Role of Hair Follicles in the Percutaneous Absorption of Caffeine. *Br. J. Clin. Pharmacol.* 65 (4), 488–492. doi:10.1111/j.1365-2125.2007.03065.x
- Ouyang, Z., and Wei, K. (2021). miRNA in Cardiac Development and Regeneration. *Cell Regen* 10 (1), 14. doi:10.1186/s13619-021-00077-5
- Palmer, J. A., Smith, A. M., Egnash, L. A., Conard, K. R., West, P. R., Burrier, R. E., et al. (2013). Establishment and Assessment of a New Human Embryonic Stem Cell-Based Biomarker Assay for Developmental Toxicity Screening. *Birth Defects Res. B* 98 (4), 343–363. doi:10.1002/bdrb.21078
- Palmer, J. A., Smith, A. M., Egnash, L. A., Colwell, M. R., Donley, E. L. R., Kirchner, F. R., et al. (2017). A Human Induced Pluripotent Stem Cell-Based *In Vitro* Assay Predicts Developmental Toxicity through a Retinoic Acid Receptor-Mediated Pathway for a Series of Related Retinoid Analogues. *Reprod. Toxicol.* 73, 350–361. doi:10.1016/j.reprotox.2017.07.011
- Parish, S. T., Aschner, M., Casey, W., Corvaro, M., Embry, M. R., Fitzpatrick, S., et al. (2020). An Evaluation Framework for New Approach Methodologies (NAMs) for Human Health Safety Assessment. *Regul. Toxicol. Pharmacol.* 112, 104592. doi:10.1016/j.yrtph.2020.104592
- Paul Friedman, K., Gagne, M., Loo, L.-H., Karamertzanis, P., Netzeva, T., Sobanski, T., et al. (2020). Utility of *In Vitro* Bioactivity as a Lower Bound Estimate of *In Vivo* Adverse Effect Levels and in Risk-Based Prioritization. *Toxicol. Sci.* 173 (1), 202–225. doi:10.1093/toxsci/kfz201
- Pessina, A., and Bonomi, A. (2007). CFU-GM Assay for Evaluation of Drug Myelotoxic Activity. *Curr. Protoc. Toxicol.* 34. 20.2.1-20.2.18. doi:10.1002/0471140856.tx2002s34
- Petzer, A., Pienaar, A., and Petzer, J. P. (2013). The Interactions of Caffeine with Monoamine Oxidase. *Life Sci.* 93 (7), 283–287. doi:10.1016/j.lfs.2013.06.020
- Piersma, A. H., Bosgra, S., van Duursen, M. B. M., Hermesen, S. A. B., Jonker, L. R. A., Kroese, E. D., et al. (2013). Evaluation of an Alternative *In Vitro* Test Battery for Detecting Reproductive Toxicants. *Reprod. Toxicol.* 38, 53–64. doi:10.1016/j.reprotox.2013.03.002
- Pohanka, M. (2015). The Perspective of Caffeine and Caffeine Derived Compounds in Therapy. *BLL* 116 (9), 520–530. doi:10.4149/bll_2015_106
- Poulsen, M. S., Rytting, E., Mose, T., and Knudsen, L. E. (2009). Modeling Placental Transport: Correlation of *In Vitro* BeWo Cell Permeability and *Ex Vivo* Human Placental Perfusion. *Toxicol. Vitro* 23 (7), 1380–1386. doi:10.1016/j.tiv.2009.07.028
- Prodromidou, K., and Matsas, R. (2019). Species-Specific miRNAs in Human Brain Development and Disease. *Front. Cel. Neurosci.* 13, 559. doi:10.3389/fncel.2019.00559
- Racz, P., Brandsma, I., Zwetsloot, T., and Hendriks, G. (2018). ReproTracker, a Human Stem Cell-Based Reporter Assay for *In Vitro* Dart Assessment. *Toxicol. Lett.* 295 (S233), S233. doi:10.1016/j.toxlet.2018.06.978
- Reynolds, J., Malcomber, S., and White, A. (2020). A Bayesian Approach for Inferring Global Points of Departure from Transcriptomics Data. *Comput. Toxicol.* 16, 100138. doi:10.1016/j.comtox.2020.100138
- Richards, D. J., Li, Y., Kerr, C. M., Yao, J., Beeson, G. C., Coyle, R. C., et al. (2020). Human Cardiac Organoids for the Modelling of Myocardial Infarction and Drug Cardiotoxicity. *Nat. Biomed. Eng.* 4 (4), 446–462. doi:10.1038/s41551-020-0539-4
- Rivkees, S. A., and Wendler, C. C. (2012). Regulation of Cardiovascular Development by Adenosine and Adenosine-Mediated Embryo protection. *ATVB* 32 (4), 851–855. doi:10.1161/ATVBAHA.111.226811
- Rivkees, S. A., Zhao, Z., Porter, G., and Turner, C. (2001). Influences of Adenosine on the Fetus and Newborn. *Mol. Genet. Metab.* 74 (1-2), 160–171. doi:10.1006/mgme.2001.3217
- Rovida, C., and Hartung, T. (2009). Re-evaluation of Animal Numbers and Costs for *In Vivo* Tests to Accomplish REACH Legislation Requirements for Chemicals - a Report by the Transatlantic Think Tank for Toxicology (T4). *ALTEX* 26 (3), 187–208. doi:10.14573/altex.2009.3.187
- Schenk, B., Weimer, M., Bremer, S., van der Burg, B., Cortvrint, R., Freyberger, A., et al. (2010). The ReProTect Feasibility Study, a Novel Comprehensive *In Vitro* Approach to Detect Reproductive Toxicants. *Reprod. Toxicol.* 30 (1), 200–218. doi:10.1016/j.reprotox.2010.05.012
- Schraml, E., Hackl, M., and Grillari, J. (2017). MicroRNAs and Toxicology: A Love Marriage. *Toxicol. Rep.* 4, 634–636. doi:10.1016/j.toxrep.2017.11.001
- Sekar, D., Mani, P., Biruntha, M., Sivagurunathan, P., and Karthigeyan, M. (2019). Dissecting the Functional Role of microRNA 21 in Osteosarcoma. *Cancer Gene Ther.* 26 (7-8), 179–182. doi:10.1038/s41417-019-0092-z
- Shahbazi, M. N., Siggia, E. D., and Zernicka-Goetz, M. (2019). Self-organization of Stem Cells into Embryos: A Window on Early Mammalian Development. *Science* 364 (6444), 948–951. doi:10.1126/science.aax0164
- Shenoy, A., and Belloch, R. H. (2014). Regulation of microRNA Function in Somatic Stem Cell Proliferation and Differentiation. *Nat. Rev. Mol. Cel Biol.* 15 (9), 565–576. doi:10.1038/nrm3854
- Silva, C. G., Métin, C., Fazeli, W., Machado, N. J., Darmopil, S., Launay, P.-S., et al. (2013). Adenosine Receptor Antagonists Including Caffeine Alter Fetal Brain Development in Mice. *Sci. Transl. Med.* 5, 197ra104. doi:10.1126/scitranslmed.3006258
- Simmons, S. O., Fan, C.-Y., and Ramabhadran, R. (2009). Cellular Stress Response Pathway System as a sentinel Ensemble in Toxicological Screening. *Toxicol. Sci.* 111 (2), 202–225. doi:10.1093/toxsci/kfp140
- Singer, J. W., Al-Fayoumi, S., Taylor, J., Velichko, S., and O'Mahony, A. (2019). Comparative Phenotypic Profiling of the JAK2 Inhibitors Ruxolitinib, Fedratinib, Momelotinib, and Pacritinib Reveals Distinct Mechanistic Signatures. *PLoS One* 14 (9), e0222944. doi:10.1371/journal.pone.0222944
- Smirnova, L., Block, K., Sittka, A., Oelgeschläger, M., Seiler, A. E. M., and Luch, A. (2014). MicroRNA Profiling as Tool for *In Vitro* Developmental Neurotoxicity Testing: the Case of Sodium Valproate. *PLoS One* 9 (6), e98892. doi:10.1371/journal.pone.0098892
- Sozen, B., Jorgensen, V., Weatherbee, B. A. T., Chen, S., Zhu, M., and Zernicka-Goetz, M. (2021). Reconstructing Aspects of Human Embryogenesis with Pluripotent Stem Cells. *Nat. Commun.* 12 (1), 5550. doi:10.1038/s41467-021-25853-4
- Spinu, N., Bal-Price, A., Cronin, M. T. D., Enoch, S. J., Madden, J. C., and Worth, A. P. (2019). Development and Analysis of an Adverse Outcome Pathway Network for Human Neurotoxicity. *Arch. Toxicol.* 93 (10), 2759–2772. doi:10.1007/s00204-019-02551-1
- Teerlink, J. R., Iragui, V. J., Mohr, J. P., Carson, P. E., Hauptman, P. J., Lovett, D. H., et al. (2012). The Safety of an Adenosine A1-Receptor Antagonist, Rolofylline, in Patients with Acute Heart Failure and Renal Impairment. *Drug Saf.* 35 (3), 233–244. doi:10.2165/11594680-000000000-00000
- Thomas, R. S., Philbert, M. A., Auerbach, S. S., Wetmore, B. A., Devito, M. J., Cote, I., et al. (2013). Incorporating New Technologies into Toxicity Testing and Risk Assessment: Moving from 21st century Vision to a Data-Driven Framework. *Toxicol. Sci.* 136 (1), 4–18. doi:10.1093/toxsci/kft178
- Thomas, R. S., Bahadori, T., Buckley, T. J., Cowden, J., Deisenroth, C., Dionisio, K. L., et al. (2019). The Next Generation Blueprint of Computational Toxicology at the U.S. Environmental Protection Agency/U.S. Environmental Protection Agency. *Toxicol. Sci.* 169 (2), 317–332. doi:10.1093/toxsci/kfz058
- Timmer, J. R., Wang, C., and Niswander, L. (2002). BMP Signaling Patterns the Dorsal and Intermediate Neural Tube via Regulation of Homeobox and helix-loop-helix Transcription Factors. *Development* 129 (10), 2459–2472. doi:10.1242/dev.129.10.2459
- Tracy, T. S., Venkataramanan, R., Glover, D. D., and Caritis, S. N. (2005). National Institute for Child, H., and Human Development Network of Maternal-Fetal-Medicine, UTemporal Changes in Drug Metabolism (CYP1A2, CYP2D6 and CYP3A Activity) during Pregnancy. *Am. J. Obstet. Gynecol.* 192 (2), 633–639. doi:10.1016/j.ajog.2004.08.030
- Tweede, S., Braschi, B., Gray, K., Jones, T. E. M., Seal, R. L., Yates, B., et al. (2021). Genenames.org: The HGNC and VGNC Resources in 2021. *Nucleic Acids Res.* 49 (D1), D939–D946. doi:10.1093/nar/gkaa980
- USDA (2015). Scientific Report of the 2015 Dietary Guidelines Advisory Committee. Available at: <https://health.gov/sites/default/files/2019-09/Scientific-Report-of-the-2015-Dietary-Guidelines-Advisory-Committee.pdf>.
- van den Akker, E., Forlani, S., Chawengsaksophak, K., de Graaff, W., Beck, F., Meyer, B. I., et al. (2002). Cdx1andCdx2have Overlapping Functions in

- Anteroposterior Patterning and Posterior axis Elongation. *Development* 129 (9), 2181–2193. doi:10.1242/dev.129.9.2181
- van der Burg, B., Wedeby, E. B., Dietrich, D. R., Jaworska, J., Mangelsdorf, I., Paune, E., et al. (2015). The ChemScreen Project to Design a Pragmatic Alternative Approach to Predict Reproductive Toxicity of Chemicals. *Reprod. Toxicol.* 55, 114–123. doi:10.1016/j.reprotox.2015.01.008
- van Gelder, M. M. H. J., van Rooij, I. A. L. M., de Jong-van den Berg, L. T. W., and Roeleveld, N. (2014). Teratogenic Mechanisms Associated with Prenatal Medication Exposure. *Therapies* 69 (1), 13–24. doi:10.2515/therapie/2014003
- Vasconcellos, R., Alvarenga, É. C., Parreira, R. C., Lima, S. S., and Resende, R. R. (2016). Exploring the Cell Signalling in Hepatocyte Differentiation. *Cell Signal.* 28 (11), 1773–1788. doi:10.1016/j.cellsig.2016.08.011
- Villeneuve, D., Volz, D. C., Embry, M. R., Ankley, G. T., Belanger, S. E., Léonard, M., et al. (2014). Investigating Alternatives to the Fish Early-life Stage Test: A Strategy for Discovering and Annotating Adverse Outcome Pathways for Early Fish Development. *Environ. Toxicol. Chem.* 33 (1), 158–169. doi:10.1002/etc.2403
- Wallace, D. R., Taalab, Y. M., Heinze, S., Tariba Lovaković, B., Pizent, A., Renieri, E., et al. (2020). Toxic-Metal-Induced Alteration in miRNA Expression Profile as a Proposed Mechanism for Disease Development. *Cells* 9 (4), 901. doi:10.3390/cells9040901
- Wang, Y., and Zhao, S. P. (2010). “Placental Blood Circulation,” in *Vascular Biology of the Placenta*. San Rafael, CA: Morgan and Claypool Publishers.
- Wetmore, B. A., Wambaugh, J. F., Ferguson, S. S., Li, L., Clewell, H. J., 3rd, Judson, R. S., et al. (2013). Relative Impact of Incorporating Pharmacokinetics on Predicting *In Vivo* hazard and Mode of Action from High-Throughput *In Vitro* Toxicity Assays. *Toxicol. Sci.* 132 (2), 327–346. doi:10.1093/toxsci/kft012
- Wiedenmann, S., Breunig, M., Merkle, J., von Toerne, C., Georgiev, T., Moussus, M., et al. (2021). Single-cell-resolved Differentiation of Human Induced Pluripotent Stem Cells into Pancreatic Duct-like Organoids on a Microwell Chip. *Nat. Biomed. Eng.* 5 (8), 897–913. doi:10.1038/s41551-021-00757-2
- Wilderman, A., VanOudenhove, J., Kron, J., Noonan, J. P., and Cotney, J. (2018). High-Resolution Epigenomic Atlas of Human Embryonic Craniofacial Development. *Cel Rep.* 23 (5), 1581–1597. doi:10.1016/j.celrep.2018.03.129
- Worringer, K. A., Rand, T. A., Hayashi, Y., Sami, S., Takahashi, K., Tanabe, K., et al. (2014). The let-7/LIN-41 Pathway Regulates Reprogramming to Human Induced Pluripotent Stem Cells by Controlling Expression of Prodifferentiation Genes. *Cell Stem Cell* 14 (1), 40–52. doi:10.1016/j.stem.2013.11.001
- Wu, S., Fisher, J., Naciff, J., Laufersweiler, M., Lester, C., Daston, G., et al. (2013). Framework for Identifying Chemicals with Structural Features Associated with the Potential to Act as Developmental or Reproductive Toxicants. *Chem. Res. Toxicol.* 26 (12), 1840–1861. doi:10.1021/tx400226u
- Xiang, L., Yin, Y., Zheng, Y., Ma, Y., Li, Y., Zhao, Z., et al. (2020). A Developmental Landscape of 3D-Cultured Human Pre-gastrulation Embryos. *Nature* 577 (7791), 537–542. doi:10.1038/s41586-019-1875-y
- Yang, H., Ma, Q., Wang, Y., and Tang, Z. (2020). Clinical Application of Exosomes and Circulating microRNAs in the Diagnosis of Pregnancy Complications and Foetal Abnormalities. *J. Transl. Med.* 18 (1), 32. doi:10.1186/s12967-020-02227-w
- Yeakley, J. M., Shepard, P. J., Goyena, D. E., VanSteenhouse, H. C., McComb, J. D., and Seligmann, B. E. (2017). A Trichostatin A Expression Signature Identified by TempO-Seq Targeted Whole Transcriptome Profiling. *PLoS One* 12 (5), e0178302. doi:10.1371/journal.pone.0178302
- Zhang, B., Kirov, S., and Snoddy, J. (2005). WebGestalt: an Integrated System for Exploring Gene Sets in Various Biological Contexts. *Nucleic Acids Res.* 33, W741–W748. doi:10.1093/nar/gki475
- Zheng, Y., Xue, X., Shao, Y., Wang, S., Esfahani, S. N., Li, Z., et al. (2019). Controlled Modelling of Human Epiblast and Amnion Development Using Stem Cells. *Nature* 573 (7774), 421–425. doi:10.1038/s41586-019-1535-2
- Zurlinden, T. J., Sali, K. S., Rush, N., Kothiya, P., Judson, R. S., Houck, K. A., et al. (2020). Profiling the ToxCast Library with a Pluripotent Human (H9) Stem Cell Line-Based Biomarker Assay for Developmental Toxicity. *Toxicol. Sci.* 174 (2), 189–209. doi:10.1093/toxsci/kfaa014

Conflict of Interest: Authors RR, MB, PC, MD, JH, HL, IM, JR, KS, WS, SS, AW, and PK was employed by the company Unilever.

Publisher’s Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright © 2022 Rajagopal, Baltazar, Carmichael, Dent, Head, Li, Muller, Reynolds, Sadh, Simpson, Spriggs, White and Kukic. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.