25th anniversary of the Berlin Workshop on Developmental Toxicology

DevTox database update, challenges in risk assessment of developmental neurotoxicity and alternative methodologies in bone development and growth

Marx-Stoelting, Philip; Solano, Marize de L M; Aoyama, Hiroaki; Adams, Ralf H; Bal-Price, Anna; Buschmann, Jochen; Chahoud, Ibrahim; Clark, Ruth; Fang, Tian; Fujiwara, Michio

Total number of authors:
30

Published in:
Reproductive Toxicology

Link to article, DOI:
[10.1016/j.reprotox.2020.11.003](https://doi.org/10.1016/j.reprotox.2020.11.003)

Publication date:
2021

Document Version
Peer reviewed version

[Link back to DTU Orbit](#)

Citation (APA):

Marx-Stoelting, P., Solano, M. D. L. M., Aoyama, H., Adams, R. H., Bal-Price, A., Buschmann, J., Chahoud, I., Clark, R., Fang, T., Fujiwara, M., Gelinsky, M., Grote, K., Horimoto, M., Bennekou, S. H., Kellner, R., Kuwagata, M., Leist, M., Lang, A., Li, W., ... Solecki, R. (2021). 25th anniversary of the Berlin Workshop on Developmental Toxicology: DevTox database update, challenges in risk assessment of developmental neurotoxicity and alternative methodologies in bone development and growth. *Reproductive Toxicology*, 100, 155-162. <https://doi.org/10.1016/j.reprotox.2020.11.003>

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Journal Pre-proof

25th anniversary of the Berlin Workshop on Developmental Toxicology: DevTox database update, challenges in risk assessment of developmental neurotoxicity and alternative methodologies in bone development and growth

Philip Marx-Stoelting, Marize de L.M. Solano, Hiroaki Aoyama, Ralf H. Adams, Anna Bal-Price, Jochen Buschmann, Ibrahim Chahoud, Ruth Clark, Tian Fang, Michio Fujiwara, Michael Gelinsky, Konstanze Grote, Masao Horimoto, Susanne Hougaard Bennekou, Rupert Kellner, Makiko Kuwagata, Marcel Leist, Annemarie Lang, Weihua Li, Alberto Mantovani, Susan L. Makris, Francisco Paumgarten, Monique Perron, Magdalini Sachana, Anne Schmitt, Steffen Schneider, Gilbert Schönfelder, Frank Schulze, Kohei Shiota, Rol Solecki

PII: S0890-6238(20)30254-9

DOI: <https://doi.org/10.1016/j.reprotox.2020.11.003>

Reference: RTX 8060

To appear in: *Reproductive Toxicology*

Received Date: 26 October 2020

Please cite this article as: Marx-Stoelting P, Solano MdLM, Aoyama H, Adams RH, Bal-Price A, Buschmann J, Chahoud I, Clark R, Fang T, Fujiwara M, Gelinsky M, Grote K, Horimoto M, Bennekou SH, Kellner R, Kuwagata M, Leist M, Lang A, Li W, Mantovani A, Makris SL, Paumgarten F, Perron M, Sachana M, Schmitt A, Schneider S, Schönfelder G, Schulze F, Shiota K, Solecki R, 25th anniversary of the Berlin Workshop on Developmental Toxicology: DevTox database update, challenges in risk assessment of developmental neurotoxicity and alternative methodologies in bone development and growth, *Reproductive Toxicology* (2020), doi: <https://doi.org/10.1016/j.reprotox.2020.11.003>

This is a PDF file of an article that has undergone enhancements after acceptance, such as the addition of a cover page and metadata, and formatting for readability, but it is not yet the definitive version of record. This version will undergo additional copyediting, typesetting and review before it is published in its final form, but we are providing this version to give early visibility of the article. Please note that, during the production process, errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

© 2020 Published by Elsevier.

**25th anniversary of the Berlin Workshop on Developmental Toxicology:
DevTox database update, challenges in risk assessment of developmental
neurotoxicity and alternative methodologies in bone development and
growth**

Philip Marx-Stoelting^{a,*}, Marize de L.M. Solano^{a,1*}, Hiroaki Aoyama^b, Ralf H. Adams^c, Anna Bal-Price^d, Jochen Buschmann^e, Ibrahim Chahoud^f, Ruth Clark^g, Tian Fangⁿ, Michio Fujiwara^h, Michael Gelinskyⁱ, Konstanze Grote^f, Masao Horimoto^v, Susanne Hougaard Bennekou^j, Rupert Kellner^k, Makiko Kuwagata^l, Marcel Leist^m, Annemarie Lang^f, Weihua Liⁿ, Alberto Mantovani^o, Susan L. Makris^p, Francisco Paumgarten^q, Monique Perron^u, Magdalini Sachana^r, Anne Schmitt^a, Steffen Schneider^s, Gilbert Schönfelder^{a,f}, Frank Schulze^a, Kohei Shiotani^t and Roland Solecki^a

^aGerman Federal Institute for Risk Assessment, Berlin, Germany

^bInstitute of Environmental Toxicology, Japan

^cMax Planck Institute for Molecular Biomedicine, Muenster, Germany

^dEuropean Commission, Joint Research Centre (JRC), Ispra, Italy

^ePrivate Consultant, Hannover, Germany

^fCharité – Universitätsmedizin Berlin, corporate member of Freie Universität Berlin, Humboldt-Universität zu Berlin, and Berlin Institute of Health

^gRuth Clark Associates Ltd., United Kingdom

^hAstellas Pharma Inc., Tokyo, Japan

ⁱTechnische Universität Dresden, Dresden, Germany

^jTechnical University of Denmark, Lyngby, Denmark

^kFraunhofer Institute for Toxicology and Experimental Medicine, Hannover, Germany,

^lNational Institute of Health Sciences, Kanagawa, Japan

^mUniversity of Konstanz, Konstanz, Germany

ⁿNHC Key Lab. of Reproduction Regulation (Shanghai Institute of Planned Parenthood Research), Fudan University, China

^oIstituto Superiore di Sanità (ISS), Rome, Italy

^pU.S. Environmental Protection Agency, Office of Research and Development, Center for Public Health and Environmental Assessment, Washington, D.C., USA

^qOswaldo Cruz Foundation, Brazil

^rOrganisation for Economic Co-operation and Development (OECD), Environment Health and Safety Division, Paris, France

^sBASF SE, Ludwigshafen, Germany

^tKyoto University, Kyoto, Japan

^uU.S. Environmental Protection Agency, Office of Pesticides Programs, Washington, D.C., USA

^v Chiba Institute of Science, Chiba, Japan

¹corresponding author: devtox@bfr.bund.de: Bundesinstitut für Risikobewertung, Max-Dohrn Strasse 8-10, D-10589 Berlin, Germany

* Contribute equally

Highlights

- The Berlin Workshops have a 25 year history of bringing together experts on developmental toxicity
- Improving categorization of developmental effects is crucial for harmonization and risk assessment
- Developing AOPs is a fundamental improvement to guide screening and testing for DNT
- More interdisciplinary research will accelerate new technologies to test developmental toxicity

Journal Pre-proof

Abstract

25 years after the first Berlin Workshop on Developmental Toxicity this 10th Berlin Workshop aimed to bring together international experts from authorities, academia and industry to consider scientific, methodologic and regulatory aspects in risk assessment of developmental toxicity and to debate alternative strategies in testing developmental effects in the future. Proposals for improvement of the categorization of developmental effects were discussed as well as the update of the DevTox database as valuable tool for harmonization.

The development of adverse outcome pathways relevant to developmental neurotoxicity (DNT) was debated as a fundamental improvement to guide the screening and testing for DNT using alternatives to animal methods.

A further focus was the implementation of an *in vitro* mechanism-based battery, which can support various regulatory applications associated with the assessment of chemicals and mixtures. More interdisciplinary and translation research should be initiated to accelerate the development of new technologies to test developmental toxicity.

Technologies in the pipeline are (i) high throughput imaging techniques, (ii) models for DNT screening tests, (iii) use of computer tomography for assessment of thoracolumbar supernumerary ribs in animal models, and (iv) 3D biofabrication of bone development and regeneration tissue models. In addition, increased collaboration with the medical community was suggested to improve the relevance of test results to humans and identify more clinically relevant endpoints.

Finally, the participants agreed that this conference facilitated better understanding innovative approaches that can be useful for the identification of developmental health risks due to exposure to chemical substances.

Key words: DevTox-Project, grey zone anomalies, neurodevelopmental toxicology testing, bone formation, alternative methods, risk assessment.

1. Introduction.

The Berlin Workshops are a series of international meetings, originally organised with the support of the International Program on Chemical Safety (IPCS), the first of which took place in 1995 [1]. The main aims of the Berlin Workshops were established during the first two meetings of the group in 1995 and 1998. A major aim was to develop and publish an international glossary of terminology [2-4].

The third and fourth workshops were set up to address the ambiguous and inconsistent terminology, often used to describe findings in developmental toxicology studies, which was leading to difficulties in making regulatory decisions [2-4]. Therefore, the second major aim was to develop a uniform international vocabulary to describe findings observed in developmental toxicology studies and to harmonize the terms used for categorization of these findings. The fifth Workshop, held in 2005, concentrated on a proposal to internationally harmonize terminology in developmental toxicology [5]. In the following workshops this work was extended and the freely accessible DevTox database was established (www.devtox.org) in line with the above mentioned aims.

At the last Berlin Workshop in 2018, the need for a harmonised terminology for classification of anomalies in laboratory animals in developmental toxicity studies in the context of human health risk assessment was addressed and the DevTox database was identified as an extremely valuable tool [6]. Participants agreed that one of the major challenges for testing developmental toxicity in the 21st century is the development of animal-free test strategies and alternatives to animal testing that can provide information relevant for humans in a rapid, efficient, and mechanistically informative manner.

The “10th Berlin Workshop on Developmental Toxicology” aimed to bring together international experts from authorities, research institutions, academia and industry (i) to discuss the final update of the DevTox database, (ii) to consider specific aspects in regulatory risk assessment of developmental neurotoxicity (DNT) and (iii) to debate alternative strategies in testing developmental effects in the future, with the main goal to improve harmonisation in the worldwide assessment of developmental findings and laboratory investigations. The program included three main sessions in which the topics were ‘The DevTox-Project’, ‘Risk Assessment – Developmental Neurotoxicity’ and ‘Future Methodology in DevTox’.

Spanning a period of 25 years, the results from the first workshop in 1995 to the tenth workshop in 2020 are summarized on the DevTox-Website¹, which now contains more than 2,900 images with examples for external, skeletal, visceral and maternal-foetal anomalies in nine different species. The historic, scientific and technical background was presented in key note lectures by the main DevTox partners: BfR, Charité-Universitätsmedizin Berlin and Fraunhofer ITEM, and these are published on the Web-site of the BfR-Academy², together with all other presentations and Live stream recordings of the workshop.

2. Update of the DevTox-Project

2.1 Follow-up activities of the Berlin Workshop and the JTS (Japanese Teratology Society) and ETS (European Teratology Society)

At the 2nd Berlin Workshop, in 1998, participants discussed the advantages and disadvantages of using a categorization system to indicate the relative importance of different findings, particularly in terms of potential harm to the organism in future life, and definitions for two categories of finding ('malformation' and 'variation') were proposed. The agreed definitions were [2]:

Name of category	Definition
Malformation	A permanent structural change that likely affect adversely the survival or health of the species under investigation.
Variation	A change that occurs within the normal population under investigation and is unlikely to adversely affect survival or health.

Following surveys in which expert examiners were asked to use the above definitions to categorize findings listed in the current international glossary [1] to one or other of these categories, it became clear that many of the terms listed in the glossary did not fit readily into either category. This led to creation of a third group of descriptive terms, originally

¹ https://www.devtox.org/workshops_en.php

² <https://www.bfr-akademie.de/index.php/english/archive/2020/10th-berlin-workshop-devtox.html>

known as 'U' ('cannot decide') and subsequently renamed 'grey zone anomalies'. In 2005, following further surveys in which participants were asked to fit the glossary terms into one of the three groups (malformation, variation or grey zone anomaly). A total of 1,041 findings were selected as grey zone anomalies whilst only 701 findings were classified either as malformation or variation. Berlin Workshop members and representatives of other international groups considered the large number of grey zone findings unsatisfactory. The main reasons for the existence of the grey zone group were identified as imprecise descriptive terms and insufficient knowledge of the postnatal consequences of the findings.

The Berlin Workshop members, representatives of other international groups and especially the Japanese Teratology Society (JTS) were concerned about the practical consequences of the large number of findings allocated to the 'grey zone' by respondents to the categorization surveys conducted following the 2nd Berlin Workshop. In order to find a possible way forward to resolve the problem, the JTS Laboratory Animal Term Project (JTS LATP) conducted a survey in which opinions from 20 facilities (12 pharmaceutical industries, 6 contract research labs, 1 chemical industry and 1 environmental laboratory) and 2 experts were sought. The JTS questionnaire was limited only to 73 findings in accordance with an established JTS rationale. This work began in 2015 and was followed up by a two-step decision-making process, which included review of the questionnaire responses and discussions on the subject of experimental animal terminology at workshops that took place at JTS meetings in 2015, 2016, and 2017.

The outcome was 38 of the 73 external grey zone anomalies were categorized as malformation, three were allocated to a suggested new category, named 'non-structural abnormality', one could not be agreed upon (either malformation or non-structural abnormality) and 31 were considered 'not applicable for foetal examination in rodent/rabbit embryo foetal development' because they were indistinguishable from changes due to foetal growth (e.g. 'large foetus', 'small foetus') or easily altered by different physiological conditions during examination (e.g. discoloured or pale foetus). The acceptance criteria for this re-categorization were a minimum 80% agreement of category by the respondents to the questionnaire or recommendation by the JTS Terminology committee. This was largely comparable to the 'index of agreement' established by Berlin Workshops in 2005, in which a 'good' index of agreement (IA) was judged to be $\geq 75\%$.

The preliminary conclusions of the decision-making process were as follows:

- Functional changes (i.e. effects occurring after completion of organogenesis and considered adverse because they indicate an alteration in functionality due to pharmaco-toxicological disturbance) should be allocated to a separate ('new') category.
- Findings (such as functional changes) not related to embryofetal developmental processes should be identified and considered separately from other grey zone anomalies.
- The definitions of malformation and variation should be refined.

2.2 Japanese proposals for update of definitions and re-categorisation of grey zone anomalies

In order to gain more understanding of the rationale applied by JTS, and in view of particular concern about the proposed new JTS category named 'not applicable for foetal examination in rodent/rabbit embryo foetal development', representatives from the Berlin Workshop and the European Teratology Society (ETS) arranged a 'pre-meeting' with JTS in Tokyo 2018. The proposal for update of definitions was presented at a workshop during the 2018 annual meeting of ETS. Here, participants agreed that further consideration of these preliminary conclusions was necessary in order to arrive at a globally acceptable proposal for reduction of the grey-zone. Involvement of the United States Society for Birth Defects Research and Prevention (BDRP), formerly known as the Teratology Society, would be necessary before final agreement on any proposed changes could be reached.

- If JTS should provide clear criteria for their decision process, ETS would be willing to consider their proposal that some terms listed in the glossary could be identified as '**recommended**' whilst others (mainly terms that are only theoretically conceivable) could be identified as 'not recommended'.
- Participants of Berlin Workshops supported a JTS proposal to include an additional category of external and visceral findings, named '**non-structural anomalies**'.
- Participants of Berlin Workshops and JTS agreed that the definition for 'malformation' could be reworded.

In the first session of the 10th Berlin Workshop, the JTS presented revised proposals, as follows

- For functional changes, a new category '**Not malformation**' should be established. Examples of descriptive terms that would fall within this category would be 'haemorrhagic' or 'discoloured'. The logic of this decision is that such findings are not strictly malformations, neither are they variations; in the meanwhile, these findings should be noted since a relationship with treatment may not be ruled out altogether.
- Permanent structural changes should be considered as a warning of possible adverse consequences to human health even when no apparent adverse effect on the health or survival of adult animals in the species under investigation exist up to now. Therefore, categorization as 'malformation' would be preferred in the case of any uncommon irreversible abnormal process.

The principles detailed in the above are those used by JTS for evaluation of embryofetal development studies. Assessment of embryofetal toxicity is considered in terms of effects on two separate developmental stages. These are 'effects on the developmental process' (i.e., organogenesis), seen as embryofetal lethality and/or fetal anomaly, and 'effects on the fetus after completion of the developmental process' (i.e., after organogenesis), seen as fetal growth retardation and/or functional changes.

In consideration of all internal discussions prior to and including the year 2020, the definitive JTS LATP proposals for grey zone reduction were:

- A new category '**Functional change**' should be established
- The revised definition for malformation should be: '**A permanent structural change resulting from an abnormal developmental process but without evidence or information about the adverse effect**'.

The JTS plan for the future includes seeking confirmation of agreement with Berlin Workshops on the definitive proposals above and creation of a list of candidate findings for the two new/revised categories. Meeting attendees suggested that the ETS and the United States Society for Birth Defects Research and Prevention (BDRP) should also be involved in this process.

To further investigate and agree on this issue, the JTS, in collaboration with Berlin Workshops and other international organisations, such as ETS and TS, should provide, as soon as possible, clear justifications for possible solutions. It was suggested that further consideration and discussion of these preliminary conclusions was necessary in order to arrive at a globally acceptable proposal for reduction of the grey-zone. JTS agreed also to

provide their criteria for anomalies which should be identified as '**recommended**' whilst others could be identified as 'not recommended' as soon as possible.

2.3 View of a developmental toxicologist from the EU on the Japanese proposal

Most regulatory decisions for chemicals and pharmaceuticals, including those related to endocrine disrupters among pesticides and biocides under review in the EU, still tend to focus on potential adversity. In discussions relating to reduction of the number of grey zone anomalies, we should keep this in mind.

Characterization of grey zone anomalies and understanding whether or not they might represent health-relevant developmental effects is definitely worthwhile. Consequently, studying appropriate adverse outcome pathways (AOPs) can help to characterize if grey zone anomalies lead to adversity: in its turn this will influence decisions on classification and thus the fate of the compound.

A good example is anogenital distance (AGD). Within the OECD prenatal developmental toxicity study guideline (OECD 414) [7] AGD measurement is an add on endpoint,, which implies that it is of interest in deciding the regulatory fate of compounds under evaluation. Observation of reduced or increased AGD would generally be considered a grey zone anomaly. Actually, the relevance of AGD depends on weight and sex. For example, reduced AGD in low weight male fetuses would be considered unlikely *per se* to adversely affect the morphology or the future health of the animal, as it may simply reflect a general growth reduction. However, reduced anogenital distance in male fetuses with normal (unaltered) body weight could indicate intrauterine endocrine disruption, which is predictive of possible reproductive impairment in later life [8, 9]. Thus, in light of our understanding of the underlying biology, this grey zone finding could have regulatory relevance for later life and be influential in decisions relating to the fate of the compound.

The JTS proposal to reduce the grey zone anomalies deserves merit because it is an effort to reduce uncertainties and it is a method based on biological consideration of the likely origin of the finding and its likely consequences.

Functional changes express adversity after organogenesis but are not restricted to growth retardation and/or fetal death.

In other examples of functional change, adversity is suspected but it remains unclear if effects on health or survival persist. unclear. Fetal edema, known as hydrops fetalis [10] in man, even when it occurs in the absence of common accompanying conditions, such as cardiovascular abnormalities or multiple malformations, may result from a number of different modes of action and it is likely to be adverse. Similarly, fetal haemorrhages are considered as a grey zone anomaly, but have been shown in human medicine to be a possible result of defective functional development of the renal cortex or an altered exchange with embryonic adnexa [11, 12].

Permanent structural changes express adversity during organogenesis. Some permanent structural changes (e.g. open neural tube) are clear malformations with obvious anatomical abnormalities and impairment of multiple functions. However, regardless of severity, all permanent structural changes reflect some degree of adversity during organogenesis. Therefore, it is both conservative and scientifically sound to categorize them as malformations.

Similarly, amongst visceral findings, changes assumed to be of minor importance because the adverse consequences are unclear could be considered as malformations in light of human data. Malpositioned blood vessels have variable severity and can be associated with complex cardiovascular abnormalities (e.g., Fallot's tetralogy) but are also a component of other clinically relevant anomalies [13]. Abnormal lung lobation can reflect abnormalities in tracheobronchial branching and might increase the risk of pneumonia or dyspnoea in later life [14]. Absent liver lobe is associated with increased risk of biliary disturbances or portal hypertension [15] and might also be suggestive of altered internal architecture of the organ.

The main conclusions were:

1. Since functional changes often lie at the border with structural change, it is suggested that the name adopted for the additional category should be '**Functional Changes**' and not '**Non-structural Functional Changes**'.
2. Permanent structural changes of apparent low severity could be considered as malformations as they could be linked, through an AOP approach, to more serious conditions. Therefore, the original definition of permanent structural change [13] could be enhanced, as follows: '**A permanent structural change resulting from an abnormal developmental process that can adversely affect, or is plausibly linked to an adverse effect on, the survival or health**'

3. Possible steps forward include use of computerized image analyses to produce quantitative values for findings such as 'small', 'large', 'dilated', and use of AOP-like analyses, based on tissue-level rather than molecular and subcellular events, to understand which grey zone anomalies are the most relevant from the standpoint of adversity.

2.4 Berlin Workshop view on a new survey for re-categorization of grey zone anomalies

A number of considerations arose in respect of these suggested new definitions:

- It was considered by Berlin Workshops that malformation (structural abnormality) and non-structural abnormality should not be in the same category. It is not useful to categorize all non-structural changes as abnormalities. Non-structural findings are better classified, as abnormality or variation, on a case by case basis.
- The exclusion of statements relating to postnatal consequence (survival or health) was difficult to accept since this is a major consideration in the context of possible adversity.
- Participants of the Berlin Workshop understand the reason for inclusion of 'induced by developmental disturbance' in the definition of abnormality but is concerned that inclusion of this statement excludes the possibility of 'deformation' (change due to mechanical force on an otherwise normal structure)
- A result of a survey is the identification of a lack of knowledge about postnatal consequences and lack of information on the severity of findings. Consequently, a high number of glossary terms was allocated to the grey zone. Additionally, the lack of knowledge on postnatal consequences is still a major problem. The problem can only be solved by substantial research, which is unlikely being initiated in the near future.
- Many of the findings allocated to the grey zone can be re-categorized as variation or abnormality by reducing the uncertainty regarding severity. Terms such as 'long', 'short', 'dilated', 'distended' and 'bent', which can apply to many different structures (e.g., limb bones and ribs at skeletal examination, blood vessels and urinary bladder at visceral examination) and thus constitute a large proportion of the findings in the grey zone, can be qualified by severity gradings e.g. as 'severe' or 'slight'.

- Findings graded as ‘severe’ would be categorized as ‘abnormality’ (‘malformation’) and those graded as ‘slight’ as ‘variation’. Inclusion of photographs to show the nature of the findings can also make the decisions easier. In these ways, relatively small enhancements to routine reporting procedures at fetal/neonatal evaluation can reduce the uncertainty for a large proportion of the findings allocated to the grey zone.

2.5 New pictures in the DevTox Database Chinese version

The Shanghai Institute of Planned Parenthood Research (SIPPR) and Fudan University, contributed new pictures to the DevTox Database. Chinese researchers produced and shared more images of developmental abnormalities with the international community [16]. In 2017, they produced new visceral images for the DevTox database. Subsequently, during the 9th Berlin Workshop in 2018, the German and Chinese groups discussed incorporation of images from the Chinese book ‘Atlas of Common Malformations in Laboratory Rabbit and Rat’, which was published by Shanghai Science and Technology Press, into the DevTox database [17]. With mutual agreement, more than 500 images from this atlas were shared with Berlin Workshops colleagues and in total 159 images (31 external, 76 skeletal, and 52 visceral images) were uploaded to the DevTox database in 2019. These images present new anomalies not previously included in the database, some pictures are fresh tissue specimens, others hand-drawn sketches of corresponding photographic images. This update of the DevTox atlas was presented via live conference from China during the meeting. In the future, it is planned that more images will be selected for upload and more pictures will be shared with all colleagues.

3. Risk assessment- alternatives methods for screening and testing developmental neurotoxicity (DNT)

3.1 Linking test systems to the prediction of DNT

Neurodevelopment begins in the first trimester of the prenatal stage with the formation of the neural tube (about week 4-6 in humans) and complex further proliferation, migration and differentiation processes of various populations of neural precursor cells (= key neurodevelopmental processes). Many new approach methods (NAM) for hazard assessment test the interference of test chemicals with key neurodevelopmental processes [18-20].

A final (externally observable) DNT effect is always linked to a toxicity endophenotype (TEP), i.e., a state of altered functionality or connectivity of a part of the nervous system [18]. Disruption of developmental processes, such as differentiation, migration and neurite growth, can be tested in simplified, cell-based methods. If one of these tests (NAM) indicates that a process important for brain development is disturbed, one can conclude that the same chemical may lead to a TEP in man, and that a positive test result therefore indicates a potential DNT hazard [21, 22].

The SysDT project of the German Ministry of Education and Research (BMBF) is an example of a recent project covering neurodevelopmental processes by NAM. An important outcome is the neural rosette formation assay (RoFA) to identify neurotoxicants affecting very early stages of neurodevelopment and to characterize their transcriptome disturbance [23]. RoFA models the early neural tube development using human (iPSC). These are used to differentiate to neural precursor cells that self-organize to rosettes. Compounds affecting this process would be expected to lead to malformations such as spina bifida or hydrocephalus. The group developed a so-called developmental potency estimation using a mathematical algorithm to translate the number of genes changed into normalized and quantitative readouts. Major challenges are the calibration of this readout scale to human pathology and the translation of test concentrations to doses/accepted exposure levels. Eventually, a set of NAM will have to be assembled in a test battery, and clear criteria on how to combine these endpoints need to be established [24].

The DNT meeting series, as well as CAAT-Europe workshops [25-29], are focusing on regulatory implementation of NAM by providing criteria for test readiness, test performance, and later validation of the NAM [18, 22].

3.2 OECD project on DNT *in vitro* battery

The Organisation for Economic Cooperation and Development (OECD) Test Guidelines Programme, provides a forum to discuss and develop internationally harmonised test guidelines and guidance documents for assessing chemical safety on human health and environment. This programme regularly receives proposals on innovative test methods for chemical safety testing to address certain regulatory relevant endpoints that, when accepted, are included in the work plan. One of these projects, already in the work plan since 2017, regards the development of guidance on application and interpretation of *in vitro* developmental neurotoxicity (DNT) assays [30]. The overall aim of this project is to

provide a flexible and tailored DNT testing battery to address different regulatory needs, identify the current suitability of assays and give guidance on data interpretation on single assay in the battery and for the battery as a whole. It will also define a tiered approach to testing and assessment of DNT modalities not related to the endocrine system, using the integrated approach to testing and assessment (IATA) conceptual framework. OECD-wide networks of national experts from scientists in government, academia and industry provide input to this project that concerns the complex endpoint on DNT.

The guidance document builds on experience gained from case studies that focus on different regulatory scenarios that may arise, depending on potential availability of *in vivo* data for a chemical (missing, negative, positive or equivocal results) and the relevant regulatory problem formulation. Under different research leaders involved in this OECD project program, the current work in progress is:"

- US DNT-NTP (National Toxicology Programme): DNT screening and prioritization of Organophosphorus flame retardants [31].
- EU-ToxRisk: DNT hazard characterization of neonicotinoid pesticides based on NAMs [30].
- BIAC (Business at OECD): *In vitro* assays and *in silico* models to screen compounds for potential DNT activity.
- EFSA (European Food Safety Authority): IATA development for DNT for hazard characterization of pesticides.

In parallel, there is ongoing experimental work on testing chemicals using the *in-vitro* testing battery, including inter-laboratory transferability of the methods and additional work on the zebrafish model that will inform particular sections in the guidance document.

The expected outcome of the testing will be BMC (benchmark concentration)-response curves for about 120 chemicals across the assays. The results will be made publicly available through a database together with the prediction models created for the individual assays and the whole battery. Therefore, it is expected that by the end of 2021 a draft guidance on an *in vitro* DNT testing battery for processing of the data and interpretation of results, and including explicit regulatory examples, will be made available for review and comment by the national coordinators of the test guidelines programme.

There are still some challenges in the area, such as the engagement of other interested agencies around the world that regulate in particular pharmaceuticals and industrial

chemicals and the agreement on the level of validation required for the in vitro battery tests. However, some efforts to address the inter-laboratory assays transferability and the in vitro battery performance analysis to provide clarity on which chemicals can be actually tested in vitro, are already underway.

3.3 Adverse Outcome Pathways (AOPs)-driven evaluation of developmental neurotoxicity induced by mixture of environmental chemicals

Authors from regulatory agencies, academia and industry from Europe, USA and Canada signed a Consensus Statement stating that a new framework based on alternative approaches needs to be adopted for assessment of chemicals with the potential to disrupt brain development [32]. Humans, including the embryo, infants and children are indisputably co-exposed to more than one chemical at a time and the current regulation does not consider the combined effects of chemicals which can be greater than toxicity induced by a single chemical, especially in the case of chemicals acting through the same/similar MoA, and possibly leading to additive or synergistic effects [20].

The application of the AOP concept for evaluation of DNT effects induced by exposure to a mixture of environmental chemicals is the scope of JRC (European Commission, Joint Research Centre) investigations. A chemical grouping approach may be based on toxicological or biological properties, for example grouping chemicals with a similar MoA or adverse outcomes defined in AOPs. Cognitive damage in children (learning and memory impairment) is the most frequent adverse outcome identified in existing AOPs [33]. The experimental design of the study consisted of four phases where, based on the evaluation of cytotoxicity for single chemicals, repeated dose treatments with single compounds or mixtures were performed and analysis of DNT-specific endpoints for synaptogenesis, neurites, and brain derived neurotrophic factor (BDNF) expression were measured in a mixture of neuronal and glial cells derived from human induced pluripotent stem cells.

The results showed that low concentrations of single chemicals become neurotoxic in mixtures, especially for the chemicals working through similar MoA. Synaptogenesis seems to be the most sensitive DNT endpoint when analysing the synergistic effect observed in mixtures after a 14-day exposure to similar MoA chemicals. Moreover, these chemicals induced DNT effects in mixtures, showing an increased number of neurons,

impairment of neurite outgrowth and synaptogenesis, and increase BDNF levels, to a certain extent reproducing autism-like cellular changes observed in the brain of autistic children [34, 35]. This approach also allowed identification of lowest observed adverse effect concentrations (LOAECs) for single chemicals and mixtures.

The results obtained suggest that the use of human iPSC-derived mixed neuronal/glia cultures applied to a battery of assays anchored to key events of an AOP network represents a valuable approach, which permits a mechanistic understanding of induced developmental neurotoxicity [36, 37]. Now the data obtained will be used for refining the existing AOPs in a more quantitative manner and, finally, in the near future, the *in vitro* assays will be included in a test battery for regulatory purposes.

3.4 Current advances of the US-EPA strategy for testing of developmental neurotoxicity

For pesticide registration in the United States, the *in vivo* DNT study is a conditionally required study. The decision to require the study applies an information-based approach that incorporates the best available information on the chemical (e.g., hazard, pharmacokinetic, mechanistic or use pattern data) to determine whether a standard guideline study, an enhanced guideline study or an alternative study should be conducted to assess potential hazard to the developing animal or, in some cases, to support a waiver for such testing. For this process, US EPA will consider the risk of a chemical allowing for consideration of both hazard and exposure.

Currently, US EPA has approximately 100 DNT studies, reviewed by the Office of Pesticides Programs, from which only 24 DNTs have been used to set points of departure for risk assessment at the time of this workshop. Of these 24 studies, nine were selected on the basis of changes in brain morphology (neuropathology and/or morphometrics) and five on the basis of behavioral changes; the remaining studies were selected on the basis of non-neurotoxicity endpoints (e.g., pup weight decrements).

In more recent years, US EPA has shifted its testing focus from the DNT guideline study to more targeted testing based on commonly accepted modes of action (MOA). In the case of organophosphate and N-methyl carbamate insecticides, acetylcholinesterase inhibition has been selected as the critical endpoint and the focus of special testing on life stage susceptibility, including the use of comparative cholinesterase assays (CCA). For

pyrethroids, in vitro studies and physiologically based pharmacokinetic (PBPK) modelling based on the known MOA, which involves interaction with voltage-gated sodium channels leading to neurotoxicity, have been used to evaluate life stage susceptibility for this class of chemicals. For chemicals toxic to the thyroid system, the Agency's testing strategy has shifted from the DNT to the Comparative Thyroid Assay (CTA) again using a MOA-driven approach.

There is also ongoing work by EPA's Office of Research and Development (ORD) as part of an international effort to develop a battery of in vitro assays and non-mammalian test species (zebrafish) for evaluating DNT. These assays focus on addressing critical neurodevelopmental processes at the cell, tissue and organism level.

4. Future methodologies in bone development *in vitro* models to elucidate mode of action

This session focused on the challenges and opportunities to better understand bone development and bone growth on developmental outcomes using new methodological approaches and non-animal alternative testing methods for risk assessment. The new insights on bone growth and the impact of immune cells and vasculogenesis on bone development are also relevant in teratology, developmental toxicity of chemical substances, and mechanisms leading to health risk.

4.1 Investigating the interaction of bone- and immune cells in osteochondral models

4.1 Investigating the interaction of bone and immune cells in a 3D in vitro fracture gap model

The bone is a specialized organ system consisting of bone-forming and -resorbing cells which are, depending on the bone location, in a close contact to the bone marrow including e.g. haematopoietic stem cells and a multitude of different immune cells. After fracture, a fracture hematoma is formed within the gap attracting immune cells and progenitor cells such as mesenchymal stromal cells (MSCs), subsequently initiating the healing cascade. Therefore, pro- inflammatory and anti-inflammatory processes in a

hypoxic microenvironment are essential steps and most often determine the success of the subsequent bone regeneration.

In order to simulate the key aspects of the initial phase of fracture healing, Pfeiffenberger et al. [38] developed an in vitro 'fracture gap' model by combining a fracture hematoma model (controlled coagulation of human peripheral blood and MSCs measuring viability of implemented cells over time, relevant gene expression patterns and release of relevant cytokines) and bone-like components (expression of osteogenic genes and demonstration of calcification via various stainings and in vitro μ CT). The results on the testing of different medications within the 3D in vitro model showed distinct processes of the initial phase of fracture healing and demonstrate the importance of including the crosstalk between immune cells and bone.

As the model uses human MSCs, one of the technical limitations is the age of the donors and their cell quality. However, the inclusion of defined characterization steps before use are essential for standardization and are routinely applied.

4.2 Vascularization of bone during development and growth

Continuing with the investigations on bone growth, there are still difficulties in establishing models to mimic the complex signalling pathways and metabolic environments of bone vascularization during development and growth [39, 40]. The molecular regulation of bone is based on different cellular systems, including the endothelial progenitor cells and endothelial cells lining the inside of a bone are active, in a crosstalk with osteoprogenitors [41]. The arteries in bone are very thin and have a specific size at the metaphysis and the diaphysis, which determine the blood flow in the long bones and define distinct metabolic environments [42].

New models, such as the generated transgenic mice with Yap1 levels controlling angiogenesis [43], have shown that the hypoxic areas are up regulated by genes identified in the knockout mice. Verteporfin inhibits the hypo activity, reduces mineralization and possibly interfere with the Yapp/T signalling. Other areas to explore, as suggested during the workshop, include investigations on stem cells and macrophages in the vicinity of vascular cells. Overall, new and already ongoing models have to consider also the impact of a hypoxic microenvironment to understand the initial phases of angiogenesis and vascularisation of bone during development and growth.

4.3 In vitro bone models, co-cultures, biomaterials and bioprinting

To mimic the full complexity of living bone is still a challenge; however, 3D printing is a fascinating technology, suitable for fabrication of 3D tissue models and even implants. The Centre for Translational Bone, Joint and Soft Tissue Research at TU Dresden are currently focusing on extrusion-based 3D printing, called 3D plotting³. The local bone cell environment is recreated with novel pasty biomaterials, suitable for 3D plotting under mild conditions, namely calcium phosphate cement, bioglass-based composite materials and a variety of biopolymer blends. By combining collagen fibril reassembly and calcium phosphate mineralisation a biomimetic composite material ("mineralised collagen") could be produced which mimics bone extracellular matrix. From this 2D membrane-like (tape), porous three-dimensional scaffolds can also be manufactured. The next step has been the inclusion of living cells (e.g. MSC - human mesenchymal stem cells) in the 3D plotting process to create directly living artificial bone tissue constructs of clinically relevant dimensions [44]. Therefore, patient specific models, such as the proliferation of the Multiple Myeloma (MM) cells in a 3D plotted bone/bone marrow model, which contains several cell types present in the bone-bone marrow interface [45], can be established. The general challenge identified by the group is how to analyse the behaviour of individual cell types in such co-culture models. All technical possibilities, such as direct and indirect co-cultures and the use of cells from different species, have their advantages and disadvantages. The more cell types involved, the more complicated it is also to include biomaterials in such model systems. Currently, co-cultures of primary human osteocytes together with either osteoblasts [46] or osteoclasts [47] are being investigated, and triple-cultures are in preparation. Such complex in vitro bone models might be relevant in the future, e.g. for osteoporosis research, as new drugs could be tested in such systems.

4.4 Research on mechanism of supernumerary rib by use of computer Tomography

The results of a research project, using Computer Tomography (CT) to investigate the mechanism of thoracolumbar supernumerary rib (TSR) induction, demonstrated that this method could also be useful in developmental investigations in animal studies. Within the scientific community, the debate relating to TSR is based on the toxicological importance of this variation within different animal incidences (from 0.07-58.10%, depending on the

³ <https://tu-dresden.de/med/mf/tfo/forschung/biofabrikation>

strain [48]) and even human populations (about 5.8% in Japanese population [49]). The limited reliable data on TSR incidence after birth and the difficulties in distinguishing chemically induced effects from spontaneous development, based only on statistically significant results, also split research opinions on the toxicological significance of TSR.

Using new technologies, i.e. CT and a chemically induced TSR animal model, it was possible to monitor morphological changes after birth and during critical windows of development in the same animal. The results showed that TSR develops within the normal range, but does not exceed the normal range after birth and it was possible to predict postnatal changes in TSR. This could lead to the use of TSR as a toxicological landmark suitable for deriving Acute Reference Doses (ARfD) in the regulatory area. However, the critical window for TSR was different in the two drugs used in the animal model: 5-FC has a shorter critical window for TSR than sodium salicylate [52]. Additionally, there was a wide variation in the spontaneous incidence of TSR between studies and animal strains, and not all chemically induced TSRs would be detected by this methodology. Furthermore, it was noted that the key to better understanding of the importance of critical windows on the mechanism of TSR induction would still depend on the toxicokinetic properties of the compound.

5. Conclusions

In recent years, efforts to reduce the size of the grey zone group have had high priorities for both Berlin Workshops and JTS. Publication of a second more detailed international glossary of terminology and revisions to the original definitions for categorization did not yet address the problem with full success. Lack of knowledge of the postnatal consequences of fetal findings remains a major problem and cannot be solved without further research. Redefining the original category definitions to exclude the statements regarding potential 'to adversely affect the survival or health of the species under investigation', as suggested by JTS, would certainly reduce the number of grey zone findings. To further investigate and agree on this issue, the JTS is providing reasonable justifications for possible solutions. It was suggested that further consideration and discussion of these preliminary conclusions will be necessary in order to arrive at a globally acceptable proposal for reduction of the grey-zone anomalies. This could also support the application of such a categorization system for indicating potential adversity (also exploiting AOP approach), assisting in regulatory decisions and providing useful

information for risk assessment. Lack of information on the severity of the finding, the other main reason for the high number of grey zone findings, may be resolved by practical enhancements to reporting procedures at fetal/neonatal evaluation.

Another challenge in developmental toxicity has been identified in the achievements and improvements of developmental neurotoxicity testing and its implementation for chemical risk assessment. Indeed, although concern exists, the guideline DNT study has so far impacted the setting of regulatory thresholds for very few compounds. The approach of using AOP networks [37, 50-52] was considered a suitable framework for evaluation of DNT effects induced by both individual or mixtures of chemicals. The various available human *in vitro* models (NAMs) when combined seem a promising alternative to identify chemicals with potential to cause learning and memory impairment in children and other neurodevelopmental disorders [20]. Therefore, they can be included in a test battery for regulatory purposes. However, the concern of potential DNT effects still exists for many un-tested compounds. Here, the best available knowledge on the chemical (hazard, pharmacokinetic, or mechanistic data) should be utilized to determine whether a standard guideline study, an enhanced guideline study, or an alternative study could be conducted to assess potential hazard to humans as regards DNT, or in some cases to support a waiver for such testing. Overall, there are many different scopes of DNT testing that could address a plethora of regulatory problem formulations. However, it is still not clear how useful more screening assays (low specificity; difficulty of quantification) are in terms of regulatory purposes. The focus of such tests has been up to now on relative ranking and identification of “alerts” to be used for prioritising the testing of chemicals, instead of finding the exact level of human hazard for risk assessment. The ongoing experimental studies are using mainly human cell tests to achieve better predictions for human DNT effects. However, the predictivity of this approach is to be further evaluated in real-life case-studies that are currently performed as part of the OECD DNT project that will inform the guidance document on application and interpretation of *in vitro* DNT assays.

Finally, participants of the 10th Berlin workshop agreed that it is important to study more cell and cellular environment to advance teratology and developmental toxicity. This is especially important to better understand toxicity of chemical substances and mechanisms leading to health risks. Great opportunities for future developmental toxicology improvement on identification of health risks by exposure to chemical substances can also be achieved, for example by using imaging technologies such as 3D, CT, or TSR.

Overall, there is a need of new technologies for developmental toxicity testing, and today there are not only *in vivo* models, but also organoids, organ-on-a-chip models and stem cell systems that can also help to understand mechanistic activities in the living organism.

The participants agreed that a result of this conference was stimulate more research& development of innovative approaches that can be useful in the identification of health risks by exposure to chemical substances and that every effort was timeless in face of the results achieved today. The Berlin Workshop is appreciated for its successful history of more than 25 years and expects to continue being the host of the most current and interesting debates in developmental toxicity, with a growing integration of multi-related areas, perspectives and goals.

Declaration of interests

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

Acknowledgments

Special thanks to the co-rapporteurs for composing the publication and for the final editing of the text. The authors would also like to thank the team of BfR Academy and especially Verena Heinrich and Wenna Xu for their excellent technical assistance in the organization of the Workshop.

This is also in loving memory of our highly esteemed colleague Dr. Martina Rauch, who was as one of the founders of DevTox, involved in the initiation of the 25th anniversary and passed away only a few months before the workshop took place.

Presentations from the 10th Berlin Workshop on Developmental Toxicology are available on the DevTox website (www.devtox.org).

DISCLAIMER: This manuscript has been reviewed by the above mentioned institutions/organisations and approved for publication. Approval does not signify that the contents reflect the views of the institutions/organisation or of OECD member countries, nor does mention of trade names or commercial products constitute endorsement or recommendation for use. The findings and conclusions in this report have not been formally disseminated by the German Federal Institute for Risk Assessment, the U.S. EPA, the OECD or other agencies listed and should not be construed to represent any agency determination or policy.

References

- [1] Wise LD, Beck SL, Beltrame D, Beyer BK, Chahoud I, Clark RL, et al. Terminology of developmental abnormalities in common laboratory mammals (version 1). *Teratology*. 1997;55:249-92.
- [2] Chahoud I, Buschmann J, Clark R, Druga A, Falke H, Faqi A, et al. Classification terms in developmental toxicology: need for harmonisation. Report of the Second Workshop on the Terminology in Developmental Toxicology Berlin, 27-28 August 1998. *Reprod Toxicol*. 1999;13:77-82.
- [3] Solecki R, Bergmann B, Burgin H, Buschmann J, Clark R, Druga A, et al. Harmonization of rat fetal external and visceral terminology and classification. Report of the Fourth Workshop on the Terminology in Developmental Toxicology, Berlin, 18-20 April 2002. *Reprod Toxicol*. 2003;17:625-37.
- [4] Solecki R, Burgin H, Buschmann J, Clark R, Duverger M, Fialkowski O, et al. Harmonisation of rat fetal skeletal terminology and classification. Report of the Third Workshop on the Terminology in Developmental Toxicology. Berlin, 14-16 September 2000. *Reprod Toxicol*. 2001;15:713-21.
- [5] Makris SL, Solomon HM, Clark R, Shiota K, Barbellion S, Buschmann J, et al. Terminology of developmental abnormalities in common laboratory mammals (version 2). *Reprod Toxicol*. 2009;28:371-434.
- [6] Solecki R, Rauch M, Gall A, Buschmann J, Kellner R, Kucheryavenko O, et al. Update of the DevTox data database for harmonized risk assessment and alternative methodologies in developmental toxicology: Report of the 9th Berlin Workshop on Developmental Toxicity. *Reprod Toxicol*. 2019;89:124-9.
- [7] OECD. Test No. 414: Prenatal Developmental Toxicity Study. OECD Guidelines for the Testing of Chemicals. Paris: OECD Publishing; 2018.
- [8] Nelson W, Liu DY, Yang Y, Zhong ZH, Wang YX, Ding YB. In utero exposure to persistent and nonpersistent endocrine-disrupting chemicals and anogenital distance. A systematic review of epidemiological studies. *Biol Reprod*. 2020;102:276-91.
- [9] Schwartz CL, Christiansen S, Vinggaard AM, Axelstad M, Hass U, Svingen T. Anogenital distance as a toxicological or clinical marker for fetal androgen action and risk for reproductive disorders. *Archives of Toxicology*. 2019;93:253-72.
- [10] Bellini C, Donarini G, Paladini D, Calevo MG, Bellini T, Ramenghi LA, et al. Etiology of non-immune hydrops fetalis: An update. *Am J Med Genet A*. 2015;167A:1082-8.
- [11] Montemarano H, Bulas DI, Rushton HG, Selby D. Bladder distention and pyelectasis in the male fetus: causes, comparisons, and contrasts. *J Ultrasound Med*. 1998;17:743-9.
- [12] Nguyen HT, Benson CB, Bromley B, Campbell JB, Chow J, Coleman B, et al. Multidisciplinary consensus on the classification of prenatal and postnatal urinary tract dilation (UTD classification system). *J Pediatr Urol*. 2014;10:982-98.
- [13] Bravo-Valenzuela NJ, Carrilho MC, Peixoto AB, Bezerra MS, Araujo Junior E. Anatomically corrected malposition of the great arteries: a challenging fetal diagnosis. *J Matern Fetal Neonatal Med*. 2019;32:3097-101.
- [14] Chassagnon G, Morel B, Carpentier E, Ducou Le Pointe H, Sirinelli D. Tracheobronchial Branching Abnormalities: Lobe-based Classification Scheme. *Radiographics*. 2016;36:358-73.
- [15] Radin DR, Colletti PM, Ralls PW, Boswell WD, Jr., Halls JM. Agenesis of the right lobe of the liver. *Radiology*. 1987;164:639-42.
- [16] Hu J, Kan H, Tian F, Xia M, Wang Y, Xiang C, et al. The international collaboration on ATLAS of developmental abnormalities of laboratory animals. *Chinese Journal of Carcinogenesis, Teratogenesis & Mutagenesis*. 2018;30:155-7.

- [17] Zhou L, Sun Z. Atlas of Common Malformations in Laboratory Rabbit and Rat. Shanghai Science and Technology Press. Shangai, China: Shanghai Science and Technology Press; 2015.
- [18] Aschner M, Ceccatelli S, Daneshian M, Fritsche E, Hasiwa N, Hartung T, et al. Reference compounds for alternative test methods to indicate developmental neurotoxicity (DNT) potential of chemicals: example lists and criteria for their selection and use. *ALTEX*. 2017;34:49-74.
- [19] Fritsche E, Crofton KM, Hernandez AF, Hougaard Bennekou S, Leist M, Bal-Price A, et al. OECD/EFSA workshop on developmental neurotoxicity (DNT): The use of non-animal test methods for regulatory purposes. *ALTEX*. 2017;34:311-5.
- [20] Fritsche E, Grandjean P, Crofton KM, Aschner M, Goldberg A, Heinonen T, et al. Consensus statement on the need for innovation, transition and implementation of developmental neurotoxicity (DNT) testing for regulatory purposes. *Toxicol Appl Pharmacol*. 2018;354:3-6.
- [21] Bal-Price A, Crofton KM, Leist M, Allen S, Arand M, Buetler T, et al. International STakeholder NETwork (ISTNET): creating a developmental neurotoxicity (DNT) testing road map for regulatory purposes. *Arch Toxicol*. 2015;89:269-87.
- [22] Bal-Price A, Hogberg HT, Crofton KM, Daneshian M, FitzGerald RE, Fritsche E, et al. Recommendation on test readiness criteria for new approach methods in toxicology: Exemplified for developmental neurotoxicity. *ALTEX*. 2018;35:306-52.
- [23] Dreser N, Madjar K, Holzer AK, Kapitza M, Scholz C, Kranaster P, et al. Development of a neural rosette formation assay (RoFA) to identify neurodevelopmental toxicants and to characterize their transcriptome disturbances. *Arch Toxicol*. 2020;94:151-71.
- [24] Shinde V, Hoelting L, Srinivasan SP, Meisig J, Meganathan K, Jagtap S, et al. Definition of transcriptome-based indices for quantitative characterization of chemically disturbed stem cell development: introduction of the STOP-Toxukn and STOP-Toxukk tests. *Arch Toxicol*. 2017;91:839-64.
- [25] Hartung T, De Vries R, Hoffmann S, Hogberg HT, Smirnova L, Tsaïoun K, et al. Toward Good In Vitro Reporting Standards. *ALTEX*. 2019;36:3-17.
- [26] Marx U, Akabane T, Andersson TB, Baker E, Beilmann M, Beken S, et al. Biology-inspired microphysiological systems to advance patient benefit and animal welfare in drug development. *ALTEX*. 2020;37:364-94.
- [27] Rovida C, Barton-Maclaren T, Benfenati E, Caloni F, Chandrasekera PC, Chesne C, et al. Internationalization of read-across as a validated new approach method (NAM) for regulatory toxicology. *ALTEX*. 2020.
- [28] Smirnova L, Harris G, Leist M, Hartung T. Cellular resilience. *ALTEX*. 2015;32:247-60.
- [29] Smirnova L, Hogberg HT, Leist M, Hartung T. Developmental neurotoxicity - challenges in the 21st century and in vitro opportunities. *ALTEX*. 2014;31:129-56.
- [30] Sachana M, Bal-Price A, Crofton KM, Bennekou SH, Shafer TJ, Behl M, et al. International Regulatory and Scientific Effort for Improved Developmental Neurotoxicity Testing. *Toxicol Sci*. 2019;167:45-57.
- [31] Behl M, Ryan K, Hsieh JH, Parham F, Shapiro AJ, Collins BJ, et al. Screening for Developmental Neurotoxicity at the National Toxicology Program: The Future Is Here. *Toxicol Sci*. 2019;167:6-14.
- [32] Bal-Price A, Pistollato F, Sachana M, Bopp SK, Munn S, Worth A. Strategies to improve the regulatory assessment of developmental neurotoxicity (DNT) using in vitro methods. *Toxicol Appl Pharmacol*. 2018;354:7-18.
- [33] Bal-Price A, Meek MEB. Adverse outcome pathways: Application to enhance mechanistic understanding of neurotoxicity. *Pharmacol Ther*. 2017;179:84-95.
- [34] Courchesne E, Mouton PR, Calhoun ME, Semendeferi K, Ahrens-Barbeau C, Hallet MJ, et al. Neuron number and size in prefrontal cortex of children with autism. *JAMA*. 2011;306:2001-10.
- [35] Nguyen HTN, Kato H, Masuda K, Yamaza H, Hirofuji Y, Sato H, et al. Impaired neurite development associated with mitochondrial dysfunction in dopaminergic neurons differentiated from exfoliated

deciduous tooth-derived pulp stem cells of children with autism spectrum disorder. *Biochem Biophys Rep.* 2018;16:24-31.

[36] Pistollato F, Bal-Prince A, Gyves EM, et al. Adverse outcome pathway- driven assessment of developmental neurotoxicity induced by chemical mixtures using human stem cell-derived neuronal/glial culture. *Environmental Health.* in press.

[37] Spinu N, Bal-Price A, Cronin MTD, Enoch SJ, Madden JC, Worth AP. Development and analysis of an adverse outcome pathway network for human neurotoxicity. *Arch Toxicol.* 2019;93:2759-72.

[38] Pfeifferberger M, Damerau A, Ponomarev I, Bucher C, Chen Y, Barnewitz D, et al. An in vitro human-based fracture gap model - Mimicking the crosstalk between bone and immune cells. *BioRxiv.* 2020.

[39] Karsenty G, Wagner EF. Reaching a genetic and molecular understanding of skeletal development. *Dev Cell.* 2002;2:389-406.

[40] Kronenberg HM. Developmental regulation of the growth plate. *Nature.* 2003;423:332-6.

[41] Kusumbe AP, Ramasamy SK, Adams RH. Coupling of angiogenesis and osteogenesis by a specific vessel subtype in bone. *Nature.* 2014;507:323-8.

[42] Ramasamy SK, Kusumbe AP, Schiller M, Zeuschner D, Bixel MG, Milia C, et al. Blood flow controls bone vascular function and osteogenesis. *Nat Commun.* 2016;7:13601.

[43] Gnimassou O, Francaux M, Deldicque L. Hippo Pathway and Skeletal Muscle Mass Regulation in Mammals: A Controversial Relationship. *Frontiers in Physiology.* 2017;8.

[44] Ahlfeld T, Cubo-Mateo N, Cometta S, Guduric V, Vater C, Bernhardt A, et al. A Novel Plasma-Based Bioink Stimulates Cell Proliferation and Differentiation in Bioprinted, Mineralized Constructs. *ACS Appl Mater Interfaces.* 2020;12:12557-72.

[45] Braham MVJ, Ahlfeld T, Akkineni AR, Minnema MC, Dhert WJA, Oner FC, et al. Endosteal and Perivascular Subniches in a 3D Bone Marrow Model for Multiple Myeloma. *Tissue Eng Part C Methods.* 2018;24:300-12.

[46] Skottke J, Gelinsky M, Bernhardt A. In Vitro Co-culture Model of Primary Human Osteoblasts and Osteocytes in Collagen Gels. *Int J Mol Sci.* 2019;20.

[47] Bernhardt A, Osterreich V, Gelinsky M. Three-Dimensional Co-culture of Primary Human Osteocytes and Mature Human Osteoclasts in Collagen Gels. *Tissue Eng Part A.* 2020;26:647-55.

[48] Kuwagata M, Sakai Y, Tanaka S, Takashima H, Katagiri R, Matsuoka T, et al. Historical control data on developmental toxicity studies in rats. *Congenit Anom (Kyoto).* 2019;59:125-31.

[49] Nakajima A, Usui A, Hosokai Y, Kawasumi Y, Abiko K, Funayama M, et al. The prevalence of morphological changes in the thoracolumbar spine on whole-spine computed tomographic images. *Insights Imaging.* 2014;5:77-83.

[50] Beilmann M, Boonen H, Czich A, Dear G, Hewitt P, Mow T, et al. Optimizing drug discovery by Investigative Toxicology: Current and future trends. *ALTEX.* 2019;36:289-313.

[51] Leist M, Ghallab A, Graepel R, Marchan R, Hassan R, Bennekou SH, et al. Adverse outcome pathways: opportunities, limitations and open questions. *Arch Toxicol.* 2017;91:3477-505.

[52] Terron A, Bal-Price A, Paini A, Monnet-Tschudi F, Bennekou SH, Members EWE, et al. An adverse outcome pathway for parkinsonian motor deficits associated with mitochondrial complex I inhibition. *Arch Toxicol.* 2018;92:41-82.