

## Editorial Board

Article 105607

[View PDF](#)

Review Article

### **Considerations for the development of guidance on dose level selection for developmental and reproductive toxicity studies**

R.W. Lewis, A.K. Andrus, J. Arroyo, S. Brescia, ... S. Marty

Article 105585

[View PDF](#)

2022 年、欧州化学品庁は生殖発生毒性 (DART) 試験における高用量の選択に関する勧告を発表し、試験の最高用量は親動物に過剰な毒性や重度の苦痛を与えずに、生殖毒性の明確な証拠を誘発することを目指すべきであることを示しました。さらに、最近の出版物では、体重増加量の 10% 減少を用量妥当性の基準として体重の 10% 減少に置き換えるべきであると提唱しています。European Centre for Ecotoxicology and Toxicology of Chemicals の専門家は、これらの最近の進展と試験結果および解釈への影響の可能性を評価し、この勧告が OECD 試験ガイドラインまたは人道的エンドポイントガイダンスと整合していないことを確認しました。さらに、DART 試験のデータ解析から、妊娠中の母体体重の 10% 減少は体重増加量の 25% 減少に相当することが示され、これは 2010 年の ILSI/HESI ワークショップでの専門家のコンセンサスとは異なります。用量選択は、他の様々な因子を考慮した生物学的アプローチに基づくべきです。明らかな毒性を引き起こし、ホメオスタシスを抑制する過剰な用量レベルは、ヒトの健康評価に関連しない影響を引き起こす可能性があるため、避けるべきです。

In 2022, the European Chemicals Agency issued advice on the selection of high dose levels for developmental and reproductive toxicity (DART) studies indicating that the highest dose tested should aim to induce clear evidence of reproductive toxicity without excessive toxicity and severe suffering in parental animals. In addition, a recent publication advocated that a 10% decrease in body weight gain should be replaced with a 10% decrease in bodyweight as a criterion for dose adequacy. Experts from the European Centre for Ecotoxicology and Toxicology of Chemicals evaluated these recent developments and their potential impact on study outcomes and interpretation and identified that the advice was not aligned with OECD test guidelines or with humane endpoints guidance. Furthermore, data analysis from DART studies indicated that a 10% decrease in maternal body weight during gestation equates to a 25% decrease in

body weight gain, which differs from the consensus of experts at a 2010 ILSI/HESI workshop. Dose selection should be based on a biological approach that considers a range of other factors. Excessive dose levels that cause frank toxicity and overwhelm homeostasis should be avoided as they can give rise to effects that are not relevant to human health assessments.

## Commentaries

### **Sensory irritation and use of the best available science in setting exposure limits: Issues raised by a scientific panel review of formaldehyde human research studies**

Katy Goyak, Stewart Holm

Article 105587

感覚刺激性および広範な曝露が認められる高生産量化学物質として、ホルムアルデヒドのヒトの健康リスクの可能性は、多くの国際的な規制機関および科学諮問機関によってレビューされてきた。EPAの有害物質規制法 (TSCA) プログラムの後援の下で、科学委員会であるヒト研究レビュー委員会は、最近、急性吸入健康影響の証拠の重み評価におけるそれらの使用に関連して、2022年統合リスク情報システム (IRIS) ホルムアルデヒド有害性評価草案に含まれる感覚刺激性研究をレビューした。この委員会は、曝露限度(例えば、試験デザインの好み;不確実性調整)を計算する目的でこれらの研究を使用することに関する一連の勧告を出した。これらの勧告は、様々な程度の科学的コンセンサスがあるトピック領域を反映している可能性があることを考慮して、この解説は、感覚刺激性に基づく国際的なホルムアルデヒド曝露限度の間の共通性と区別を反映している。特に、科学の評価を担当する各レビュー委員会は、曝露期間またはヒトの変動性を考慮するための調整は必要ないと勧告した。これらの科学的コンセンサスの領域は、TSCAで予定されているホルムアルデヒドのリスク評価において、曝露限度を設定する目的で利用可能な最良の科学として考慮されるべきである。

As a high production volume chemical with recognized sensory irritation and widespread exposure, the human health risk potential of formaldehyde has been reviewed by many international regulatory agencies and scientific advisory bodies. A scientific panel, the Human Studies Review Board, under the auspices of the EPA's Toxic Substances Control Act (TSCA) program recently reviewed the sensory irritation studies included in the 2022 Draft Integrated Risk Information System (IRIS) Formaldehyde Hazard Assessment in the context of their use in a weight of evidence evaluation of acute inhalation health effects. This panel issued a series of recommendations on the use of these studies for the purposes of calculating exposure limits (e.g., study design preferences; uncertainty adjustment). Considering that these recommendations might reflect topic areas with varying degrees of scientific consensus, this commentary reflects

on commonalities and distinctions amongst international formaldehyde exposure limits based on sensory irritation. Notably, each review panel charged with an assessment of the science recommended that no adjustment was needed to account for either exposure duration or human variability. These areas of scientific consensus should be considered as the best available science for the purposes of setting exposure limits in the anticipated TSCA Risk Evaluation on formaldehyde.

## Discussion

### **Use of guinea pig data to obtain starting points for skin sensitisation risk assessment - A commentary**

David Basketter, Namali Corea, Marco Corvaro, Arthur Grivel, ... Christiane Wiemann  
Article 105584

植物防疫製品 (PPP) による作業員や一般公衆の皮膚感作の可能性を理解しようとする動きが高まっており、その結果、これらの製品とそのばく露シナリオに適用可能な定量的リスク評価 (QRA) 方法論を確立するための最近の構想が生まれている。皮膚感作性物質の効果的な評価には、毒性学的有害性の同定だけでなく、相対的な感作性の強さの決定も必要である。典型的には、これは局所リンパ節アッセイ (LLNA) 用量反応データの解釈によって達成され、EC3 値として知られるものを提供する。これにより、皮膚感作性物質を規定された強さのサブカテゴリーに規制上区分することが可能になったが、より重要なことに、QRA の出発点としての予想外感作誘導レベル (NESIL) の導出が可能になった。しかし、既存の物質の多くには LLNA データがなく、モルモットの古い結果のみが存在する。追加の (in vivo) 試験を回避するために、モルモットのデータと、モルモットに基づく NESIL を提供する強さのサブカテゴリーの決定に関する既存の規制ガイドラインを採用するアプローチが概説されている。このアプローチは、LLNA NESIL ベンチマークからの保守的な外挿を採用し、他の産業で既に成功している QRA プロセスのタイプの基礎として出発点を提供する。

The increasing drive to understand the likelihood of skin sensitisation from plant protection products (PPPs) in workers and the general public has resulted in recent initiatives to establish a quantitative risk assessment (QRA) methodology applicable to these products and their exposure scenarios. The effective evaluation of skin sensitising substances requires not only the identification of that toxicological hazard, but also determination of relative sensitising potency. Typically, this has been achieved by interpretation of local lymph node assay (LLNA) dose response data, delivering what is known as the EC3 value. This permitted regulatory division of skin sensitisers into defined potency sub-categories, but more importantly enabled derivation of a no expected sensitisation induction level (NESIL) as the point of departure for QRA. However, for many existing substances there is no LLNA data, only older guinea pig results exist. To

avoid additional (in vivo) testing, an approach has been outlined to employ guinea pig data and existing regulatory guidelines on the determination of potency sub-categorisation to provide a guinea pig based NESIL. The approach adopts a conservative extrapolation from LLNA NESIL benchmarks to deliver points of departure as the basis for the type of QRA process already in successful use by other industries.

## Research article

### **Detailed aggregate exposure analysis shows that exposure to fragrance ingredients in consumer products is low: Many orders of magnitude below thresholds of concern**

Isabelle Lee, Cesar Scrochi, Olive Chon, Mary Ann Cancellieri, ... Anne Marie Api

Article 105569

[View PDF](#)

Research Institute for Fragrance Materials (RIFM) と Creme Global Cremeglobal.com は共同で、香料成分の総曝露モデルを開発した。このモデルは、集団全体の個人に対する香料成分の総曝露の現実的な推定値を提供する。毒性懸念閾値 (TTC) と皮膚感作閾値 (DST) を用いて、香料物質への低曝露の程度を示した。RIFM のインベントリーに含まれる約 3000 種類の香料成分について、慢性全身、吸入、皮膚の 95 パーセンタイル曝露の合計を、それぞれの TTC または DST と比較した。さらに、代表的な香料成分を無作為に選択し、製品タイプ(すなわち、美容/パーソナルケア、家庭ケア、オーラルケア、およびエアケア)と曝露経路による曝露分布を分析した。その結果、95 パーセンタイル全身曝露と比較した場合、香料成分の 76% がそれぞれの TTC 限界値を下回っており、99% が吸入による TTC 限界値を下回っていた。製品タイプ別の 95 パーセンタイル総曝露量が最も低かったのは、家庭用ケア製品、次いでエアケア製品、オーラルケア製品だった。最も高かったのはパーソナルケア/化粧品製品だった。ほとんどの香料成分 (63%) の使用量は 1 トン未満であり、香料成分への環境曝露は低いと推定されている。

The Research Institute for Fragrance Materials (RIFM) and Creme Global [Cremeglobal.com](https://www.cremeglobal.com) partnered to develop an aggregate exposure model for fragrance ingredients. The model provides a realistic estimate of the total exposure of fragrance ingredients to individuals across a population. The Threshold of Toxicological Concern (TTC) and Dermal Sensitization Threshold (DST) were used to demonstrate the magnitude of low exposure to fragrance materials. The total chronic systemic, inhalation, and dermal 95th percentile exposures on approximately 3000 fragrance ingredients in RIFM's inventory were compared to their respective TTC or DST. Additionally, representative fragrance ingredients were randomly selected and analyzed for exposure distribution by product type (i.e., cosmetic/personal care, household care, oral care, and

air care) and route of exposure. It was found that 76 % of fragrance ingredients fall below their respective TTC limits when compared to 95th percentile systemic exposure, while 99 % are below inhalation TTC limits. The lowest 95th percentile aggregate exposure by product type was from household care products, then air care, and oral care products. The highest exposure was from personal care/cosmetic products. The volume of use for most fragrance ingredients (63 %) was <1 metric ton, estimating that environmental exposure to fragrance ingredients is likely low.

## Research article

### **Is schedule IV suvorexant an appropriate reference drug of abuse to use in preclinical abuse liability testing in the rat?**

Greet B.A. Teuns, Michela Tessari

Article 105570

[View PDF](#)

乱用に関与する既知の受容体に対する特異性が低い新規 CNS 活性薬物候補の乱用可能性は、適切な乱用の参照薬物と比較して前臨床試験が複雑である可能性がある。Schedule IV 二重オレキシン受容体拮抗薬であるスボレキサントは、Drug Discrimination Learning (DDL) 試験における参照薬物としての可能性を検討した。最初に、スボレキサントのトキシコキネティクス特性を、MC および PEG 325 の 160 mg/kg を単回経口投与した雌雄ラットで測定した。その後、325 mg/kg のスボレキサント対賦形剤の主観的影響を DDL パラダイムで評価し、血漿曝露を測定した。325 mg/kg のスボレキサントの単回投与後の雄ラットにおける平均最大血漿曝露量は、40 mg q.d. の治療量以上の用量でのヒト曝露量の 2.5 倍 (MC) から 10.5 倍 (Cmax: 1.1  $\mu$ M)、承認された最大ヒト有効用量 (20 mg, 1 日 3 回, 0.557  $\mu$ M) の 4.9 倍 (MC) から 20.8 倍 (PEG) であった。しかし、DDL 研究における 325 mg/kg の雄ラットの訓練は、それぞれの賦形剤に対する識別刺激汎化を生じなかった。スケジュール IV 二重オレキシン受容体拮抗薬であるスボレキサントは、適切な曝露にもかかわらず、ラットの DDL パラダイムにおける乱用のロバストな参照薬物として機能しなかった。

The abuse potential of novel CNS-active drug candidates with low specificity for known receptors involved in abuse might be complex to test preclinically relative to an appropriate reference drug of abuse. Suvorexant, a Schedule IV dual orexin receptor antagonist was investigated for its potential use as a reference drug in Drug Discrimination Learning (DDL) studies. Firstly, toxicokinetic properties of suvorexant were determined in male and female rats after single oral doses of 160 and 325 mg/kg in MC and PEG400. Thereafter the subjective effects of suvorexant at 325 mg/kg versus vehicle were evaluated in a DDL paradigm and plasma exposures were measured. Mean maximum plasma exposures in male rats after a single dose of 325 mg/kg suvorexant



were 2.5- (MC) to 10.5-fold (PEG400) the human exposure at supratherapeutic doses of 40 mg q.d. ( $C_{max}$ :1.1  $\mu$ M), and 4.9- (MC) to 20.8-fold (PEG400) the approved maximum human efficacious dose (20 mg q.d.; 0.557  $\mu$ M). Training male rats at 325 mg/kg in the DDL study however did not result in discriminative stimulus generalisation versus respective vehicles. Suvorexant, a Schedule IV dual orexin receptor antagonist failed to serve as a robust reference drug of abuse in the DDL paradigm in rats despite appropriate exposures.

## Research article

### **Derivation of no significant risk levels for three lower acrylates: Conclusions and recommendations from an expert panel**

C.R. Kirman, P.J. Boogaard, J.S. Bus, V.L. Dellarco, ... S.M. Hays

Article 105567

[View PDF](#)

毒性学、作用機序 (MOA)、がんリスク評価の専門家からなる委員会が、入手可能な最良の科学、データ、方法を用いて、アクリル酸メチル (MA)、アクリル酸エチル (EA)、アクリル酸 2-エチルヘキシル (2EHA) の 3 種の低級アクリル酸エステルの非有意リスクレベル (NSRLs) を導出するために取り組んだ。このレビューは、デルファイ法を修正した 5 ラウンド形式で構成された。これは、委員会メンバーから独立した熟議的な意見を収集するための体系的なプロセスであり、バイアスや集団思考の潜在的な原因を減らすためのいくつかの手続き的要素を含んでいる。用量反応評価の重要な決定に関する委員会からの意見は、MA、EA、2EHA の NSRL 値をそれぞれ 530  $\mu$ g/日 (330–800  $\mu$ g/日)、640  $\mu$ g/日 (280–670  $\mu$ g/日)、1700  $\mu$ g/日 (1300–2700  $\mu$ g/日) とした。このアプローチの新規な点は、実験用げっ歯類で腫瘍が報告されている接触点で報告された非腫瘍性病変を使用し、委員会の推奨に基づいて低用量への非線形外挿 (不確実係数アプローチ) を行ったことである。これらの値の信頼度は、試験された曝露経路(MA および EA は吸入、2EHA は経皮)に適用された曝露については中程度から高いと考えられるが、他の曝露経路に適用された場合には信頼度は低いと考えられる。

A panel of toxicology, mode of action (MOA), and cancer risk assessment experts was engaged to derive no-significant-risk-levels (NSRLs) for three lower acrylates: methyl acrylate (MA), ethyl acrylate (EA), and 2-ethylhexyl acrylate (2EHA) using the best available science, data, and methods. The review was structured as a five-round, modified Delphi format, a systematic process for collecting independent and deliberative input from panel members, and it included several procedural elements to reduce potential sources of bias and groupthink. Input from the panel for key decisions in the dose-response assessments resulted in NSRL values of 530  $\mu$ g/day (330–800  $\mu$ g/day), 640  $\mu$ g/day (280–670  $\mu$ g/day), and 1700  $\mu$ g/day (1300–2700  $\mu$ g/day) for MA, EA, and 2EHA,

respectively. Novel to this approach were the use of nonneoplastic lesions reported at point of contact where tumors have been reported in laboratory rodents, along with nonlinear extrapolation to low doses (uncertainty factor approach) based upon panel recommendations. Confidence in these values is considered medium to high for exposures applied to the routes of exposure tested (inhalation for MA and EA, dermal for 2EHA), but confidence is considered lower when applied to other routes of exposure.

## Research article

### **Chronic aquatic toxicity assessment of diverse chemicals on *Daphnia magna* using QSAR and chemical read-across**

Ankur Kumar, Vinay Kumar, Probir Kumar Ojha, Kunal Roy

Article 105572

ここでは、QSAR と化学物質読み取り法を用いて、エンドポイントとして pNOEC (mM で表した無影響濃度の負の対数) と pEC<sub>50</sub> (mM で表した半値有効濃度の負の対数) を用いて、ミジンコの慢性毒性をモデル化した。QSAR モデルは OECD ガイドラインに厳密に準拠して開発され、信頼性、予測性、正確性、ロバスト性が認められた。開発されたモデルの選択された特徴から、親油性と飽和度の増加、求電子性または電気陰性または重原子の存在、硫黄、アミンおよびそれらに関連する官能性の存在、平均原子分極率の増加、および (チオ) カルバメート (芳香族) 基の数の増加が慢性毒性の原因であることがわかった。したがって、この情報は、環境に優しく安全な化学物質の開発とデータギャップの埋め、ならびに水生生態系に慢性毒性影響を及ぼす特定された毒性化学物質の使用の削減に役立つ可能性がある。DrugBank データベースからの承認された薬物クラスおよび化学物質および製品カテゴリー (CPDat) データベースからの化学物質の多様なグループも、開発されたモデルによって評価された。

We have modeled here chronic *Daphnia* toxicity taking pNOEC (negative logarithm of no observed effect concentration in mM) and pEC<sub>50</sub> (negative logarithm of half-maximal effective concentration in mM) as endpoints using QSAR and chemical read-across approaches. The QSAR models were developed by strictly obeying the OECD guidelines and were found to be reliable, predictive, accurate, and robust. From the selected features in the developed models, we have found that an increase in lipophilicity and saturation, the presence of electrophilic or electronegative or heavy atoms, the presence of sulphur, amine, and their related functionality, an increase in mean atomic polarizability, and higher number of (thio-) carbamates (aromatic) groups are responsible for chronic toxicity. Therefore, this information might be useful for the development of environmentally friendly and safer chemicals and data-gap filling as well as reducing the use of identified toxic chemicals which have chronic toxic effects on aquatic

ecosystems. Approved classes of drugs from DrugBank databases and diverse groups of chemicals from the Chemical and Product Categories (CPDat) database were also assessed through the developed models.

## Research article

### Adverse effects in traditional and alternative toxicity tests

Patience Browne, Katie Paul Friedman, Kim Boekelheide, Russell S. Thomas

Article 105579

[View PDF](#)

化学物質安全性評価は、1 つ以上の測定されたエンドポイントを変化させる化学物質の最低レベルを定義することから始まる。この臨界影響レベルは、不確実性を説明する因子とともに、ヒトへの曝露の限界を導出するために用いられる。影響を受ける特定のメカニズムや生物学的経路に関するデータがない場合、体重や非標的臓器重量の変化などの非特異的エンドポイントが臨界影響レベルの設定に用いられる。生殖機能の障害や神経発達の変化などの特定の最終エンドポイントも、化学物質の安全限界を設定するために用いられてきた。しかし、特定の最終影響について設計された試験ガイドラインでは、同時に測定された非特異的エンドポイントは、特定のエンドポイントと同等かそれ以上に感度が高い可能性がある。これは、特定の毒性学的反応を予測するのではなく、同じ変化がヒトでも観察されると仮定せずに、保護的な臨界影響レベルを設定するために動物データがしばしば用いられることを意味する。本論文は、化学物質の有害作用の解釈方法の再考を促すことを目的としている。in vivo 毒性試験データからの非特異的エンドポイントは、広く保護的であるが必ずしも標的特異的ではない推奨曝露レベルを作成するための安全性評価係数を用いる出発点を導出するためにしばしば用いられる。

Chemical safety assessment begins with defining the lowest level of chemical that alters one or more measured endpoints. This critical effect level, along with factors to account for uncertainty, is used to derive limits for human exposure. In the absence of data regarding the specific mechanisms or biological pathways affected, non-specific endpoints such as body weight and non-target organ weight changes are used to set critical effect levels. Specific apical endpoints such as impaired reproductive function or altered neurodevelopment have also been used to set chemical safety limits; however, in test guidelines designed for specific apical effect(s), concurrently measured non-specific endpoints may be equally or more sensitive than specific endpoints. This means that rather than predicting a specific toxicological response, animal data are often used to develop protective critical effect levels, without assuming the same change would be observed in humans. This manuscript is intended to encourage a rethinking of how adverse chemical effects are interpreted: non-specific endpoints from *in vivo* toxicological



studies data are often used to derive points of departure for use with safety assessment factors to create recommended exposure levels that are broadly protective but not necessarily target-specific.

## Research article

### **Pre-clinical safety assessment of biotechnologically produced lacto-*N*-tetraose (LNT)**

Hester van der Woude, Sylvia M.J.G. Pelgrom, Carin Buskens, Roy Hoffmans, ... Dianne J. Delsing

Article 105580

ラクト-N-テトラオース (LNT) はヒト乳中に存在するオリゴ糖であり、授乳時期にもよるが、母乳中平均濃度は 0.74 から 1.07 g/L である。本研究では、*Escherichia coli* K-12 E2083 生産株により生産された LNT の前臨床安全性を評価した。LNT は細菌を用いた復帰突然変異試験および *in vitro* 小核試験のいずれにおいても陰性であり、本物質の遺伝毒性はないことが示された。OECD 408 ガイドラインに準拠したラット 90 日間経口毒性試験において、LNT は試験された最高用量までのいずれの投与群においても有害作用を誘発せず、LOAEL は決定できなかった。したがって、無毒性量 (NOAEL) は試験された最高用量、すなわち飼料中濃度 5% (w/w) に設定され、雄では  $\geq 2856$  mg/kg 体重/日、雌では  $\geq 3253$  mg/kg 体重/日に相当する。これは試験された用量レベルの範囲によって引き起こされた NOAEL の過小評価である可能性がある。本試験で得られた結果は、バイオテクノロジーで製造された他の LNT バッチを用いて得られたデータとよく一致しており、食品原料としての LNT の安全な使用を支持する。

Lacto-*N*-tetraose (LNT) is a human milk oligosaccharide with average concentrations ranging from 0.74 to 1.07 g/L in breastmilk, depending on the lactation stage. In this study, the preclinical safety of LNT produced by the *Escherichia coli* K-12 E2083 production strain was assessed. LNT was negative in both the bacterial reverse mutation assay and the *in vitro* micronucleus assay, demonstrating the absence of genotoxic potential for this substance. In the OECD 408 guideline compliant 90-day oral toxicity study rat, LNT did not induce any adverse effects in any treatment group up to and including the highest dose tested, and no LOAEL could be determined. Therefore, the no-observed-adverse effect level (NOAEL) is set at the highest dose level tested, i.e. a dietary level of 5 % (w/w), corresponding to  $\geq 2856$  mg/kg bw/day and  $\geq 3253$  mg/kg bw/day for males and females, respectively. This might be an underestimation of the NOAEL, caused by the range of dose levels tested. The results obtained in the current study are in good agreement with available data generated using other biotechnologically produced LNT batches and therefore support its safe use as a food ingredient.

## Research article

### ***In vivo* alkaline comet assay: Statistical considerations on historical negative and positive control data**

Timur Tug, Julia C. Duda, Max Menssen, Shannon Wilson Bruce, ... Christina Ziemann

Article 105583

[View PDF](#)

アルカリコメットアッセイは、異なる試験項目の DNA 損傷の可能性を特徴付けるために、異なる規制環境下で *in vivo* フォローアップ試験として頻繁に使用される。対応する OECD テストガイドライン 489 は、統計解析とヒストリカルコントロールデータ (HCD) の重要性を強調しているが、詳細な手順を提供していない。そのため、ドイツ語圏環境突然変異研究協会 (GUM) のワーキンググループ「統計」は、五つの研究所と 200 を超えるコメットアッセイ研究から HCD を収集し、いくつかの統計解析を実施した。主な結果は、(I) 観察された大きな研究所間効果は、絶対品質閾値の使用に反対である、(II) スライド上の 50%を超えるゼロ値は、スライドまたは動物の要約統計に影響を与えるため、問題があると考えられる、(III) 単一細胞データの要約尺度の種類(例えば、中央値、算術平均および幾何平均)は、HCD で得られる動物の尾の強度と試験結果に極端な差をもたらす可能性がある、であった。これらの要約値は、統計モデルの仮定によりよく適合することで解析結果の信頼性を高めるが、情報の損失を犠牲にしている。さらに、データセットにおける陰性対照群と陽性対照群の間の関係は、比、差および分位分析に基づいて、常に満足できる (または十分に) ものであった。

The alkaline comet assay is frequently used as *in vivo* follow-up test within different regulatory environments to characterize the DNA-damaging potential of different test items. The corresponding OECD Test guideline 489 highlights the importance of statistical analyses and historical control data (HCD) but does not provide detailed procedures. Therefore, the working group “Statistics” of the German-speaking Society for Environmental Mutation Research (GUM) collected HCD from five laboratories and >200 comet assay studies and performed several statistical analyses. Key results included that (I) observed large inter-laboratory effects argue against the use of absolute quality thresholds, (II) > 50% zero values on a slide are considered problematic, due to their influence on slide or animal summary statistics, (III) the type of summarizing measure for single-cell data (e.g., median, arithmetic and geometric mean) may lead to extreme differences in resulting animal tail intensities and study outcome in the HCD. These summarizing values increase the reliability of analysis results by better meeting statistical model assumptions, but at the cost of information loss. Furthermore, the relation between negative and positive control groups in the data set was always satisfactorily (or sufficiently) based on ratio, difference and quantile analyses.

## Research article

### **Repeat dose and reproductive toxicity of thrombopoietin mimic peptide in Sprague-Dawley rats**

Shidong Qiu, Yuji Liang, Xiaobo Wang, Xiue Li, ... Yanling Mu

Article 105581

トロンボポエチン受容体作動薬であるトロンボポエチン模倣ペプチド (TMP) の毒性について検討した。本研究では、TMP の反復投与毒性と生殖発生毒性試験(セグメント I、II、III)における毒性を評価した。Sprague-Dawley (SD) ラットに TMP を 5, 15, 50 mcg/kg 皮下投与した。反復投与毒性試験では、週三回 26 週間投与し、4 週間の回復期間を設けた。TMP は抗薬物抗体を産生し、血小板数増加、巨核球増殖を誘導した。一方、血小板数は徐々に減少し、雄ラットでは 4 週間後に正常化した。その他、骨髓の骨髓線維化、肝臓の髄外造血、脾臓のリンパ球減少、骨過形成が認められた。投与に関連した全ての影響は回復後に回復した。雌ラットにおける反復投与毒性の NOAEL は 5 mcg/kg であった。生殖/発生毒性(セグメント I, III)では、死亡例は認められず、一般毒性及び生殖機能の異常は認められなかった。胚・胎児発生毒性試験 (セグメント II) では、50 mcg/kg 群で吸収胚数の有意な増加が認められた。これらのラットにおける生殖/発生毒性に関する NOAEL は 15 mcg/kg であった。

Thrombopoietin mimic peptide (TMP) is a novel thrombopoietin receptor agonist. In this report, we evaluated the potential toxicity of TMP in repeat-dose toxicity and reproductive/developmental toxicity studies (segment I, II, III). TMP was administered subcutaneously to Sprague-Dawley (SD) rats at 5, 15 or 50 mcg/kg. In repeat-dose toxicity study, the rats were administrated three times a week for 26 week with a 4-week recovery. TMP could produce anti-drug antibodies and induce platelet counts increase, megakaryocyte proliferation. While platelet counts decreased gradually and returned to normal after 4 weeks in male rats. Other significant findings included myelofibrosis of bone marrow, hepatic extramedullary hematopoiesis, splenic lymphocytic depletion and bone hyperostosis. All treatment-related effects were reversed following recovery. The NOAEL of repeat-dose toxicity in female rats is 5 mcg/kg. In the reproductive/developmental toxicity (segment I, III), no deaths occurred, and no general toxicological effects or abnormal reproductive functions were observed. In embryo-fetal developmental toxicity study (segment II), the number of resorbed fetuses in the 50 mcg/kg group was significantly increased. The NOAEL as related to reproductive/developmental toxicity in these rats was 15 mcg/kg.

## Research article

### Replacing concurrent controls with virtual control groups in rat toxicity studies

Alexander Gurjanov, Carlos Vieira-Vieira, Julia Vienenkoetter, Lea A.I. Vaas, Thomas Steger-Hartmann

Article 105592

[View PDF](#)

非臨床毒性における仮想対照群 (VCG) は、同時対照群の動物を置き換えるために適切な過去の対照データを使用するという概念を表す。標準化された研究から収集された過去の対照データは、VCG を構築するための基礎として役立ち、レガシー研究報告書は VCG の性能を評価するためのベンチマークとして使用できる。レガシー研究の同時対照を VCG に置き換えることは、これらの研究の結果を理想的に再現するはずである。毒性所見の程度が異なる 3 つの 4 週間のラット経口毒性レガシー研究に基づいて、我々は異なるレベルで VCG の性能を評価する概念を開発した: (i) 同時対照からの統計的に有意な偏差を再現する、(ii) 被験物質に関連した影響を再現する、(iii) 閾値用量、標的臓器、毒性学的バイオマーカー (臨床病理学) および可逆性の観点から毒性研究の結論を再現する VCG の能力。VCG は統計的結果を再現する能力が低~中程度であることを示したが、一般的な研究結論は変わらなかった。我々の結果は、慎重に選択した過去の対照データを、一般的な研究結論を損なうことなく同時対照に置き換えるために使用できることを初めて示唆している。さらに、開発された手順とワークフローは、規制毒性学で使用するための仮想対照の将来の検証の基礎を築く。

Virtual control groups (VCGs) in nonclinical toxicity represent the concept of using appropriate historical control data for replacing concurrent control group animals. Historical control data collected from standardized studies can serve as base for constructing VCGs and legacy study reports can be used as a benchmark to evaluate the VCG performance. Replacing concurrent controls of legacy studies with VCGs should ideally reproduce the results of these studies. Based on three four-week rat oral toxicity legacy studies with varying degrees of toxicity findings we developed a concept to evaluate VCG performance on different levels: the ability of VCGs to (i) reproduce statistically significant deviations from the concurrent control, (ii) reproduce test substance-related effects, and (iii) reproduce the conclusion of the toxicity study in terms of threshold dose, target organs, toxicological biomarkers (clinical pathology) and reversibility. Although VCGs have shown a low to moderate ability to reproduce statistical results, the general study conclusions remained unchanged. Our results provide a first indication that carefully selected historical control data can be used to replace concurrent control without impairing the general study conclusion. Additionally, the developed procedures and workflows lay the foundation for the future validation of

virtual controls for a use in regulatory toxicology.

## Research article

### **Genotoxicity evaluation of food additive titanium dioxide using a battery of standard *in vivo* tests**

Chunlai Liang, Xiaopeng Zhang, Jin Fang, Nana Sun, ... Xudong Jia

Article 105586

二酸化チタンナノ粒子 (NP) の使用の増加は、食品添加物 TiO<sub>2</sub> の安全性に関する懸念を引き起こしている。TiO<sub>2</sub> は、遺伝毒性に関する懸念のため、EFSA によってもはや安全ではないと考えられているが、食品添加物としての TiO<sub>2</sub> の安全性については意見が対立しており、食品添加物 TiO<sub>2</sub> について実施された *in vivo* 遺伝毒性試験の数は限られている。食品添加物 TiO<sub>2</sub> の潜在的な遺伝毒性を調べるために、哺乳類赤血球小核試験、哺乳類骨髄染色体異常試験および *in vivo* 哺乳類アルカリコメットテストを含む一連の標準的な *in vivo* 試験を用いて、市販の食品添加物 TiO<sub>2</sub> (平均サイズは 135.54 ± 41.01 nm、範囲は 60.83 から 230.16 nm、NP は数で 30% を占める) の遺伝毒性を評価した。250, 500 および 1000 mg/kgBW の用量で 15 日間連続して胃内投与した後、食品添加物 TiO<sub>2</sub> は、マウスで骨髄小核または染色体異常の頻度を増加させず、ラット肝細胞で DNA 鎖切断を誘導しなかった。これらの結果は、本研究の条件下で、食品添加物 TiO<sub>2</sub> は NP の画分を含むが、遺伝毒性の可能性を持たないことを示す。

The increasing use of titanium dioxide (TiO<sub>2</sub>) nanoparticles (NPs) has raised concern about the safety of food additive TiO<sub>2</sub>. TiO<sub>2</sub> has been considered no longer safe by EFSA due to concerns over genotoxicity, however, there are conflicting opinions upon the safety of TiO<sub>2</sub> as a food additive, and the number of *in vivo* genotoxicity studies conducted on food additive TiO<sub>2</sub> was limited. In order to investigate the potential genotoxicity of food additive TiO<sub>2</sub>, we evaluated the genotoxicity of a commercial food additive TiO<sub>2</sub> (average size of 135.54 ± 41.01 nm, range from 60.83 to 230.16 nm, NPs account for 30% by number) using a battery of standard *in vivo* tests, including mammalian erythrocyte micronucleus test, mammalian bone marrow chromosomal aberration test and *in vivo* mammalian alkaline comet test. After 15 days of consecutive intragastric administration at doses of 250, 500, and 1000 mg/kgBW, food additive TiO<sub>2</sub> neither increased the frequencies of bone marrow micronuclei or chromosomal aberration in mice, nor induced DNA strand breakage in rat liver cells. These results indicate that under the condition of this study, food additive TiO<sub>2</sub> does not have genotoxic potential although it contains a fraction of NPs.



## Research article

### The role of trust in the use of artificial intelligence for chemical risk assessment

Pim N.H. Wassenaar, Jordi Minnema, Jelle Vriend, Willie J.G.M. Peijnenburg, ... Anne Kienhuis

Article 105589

[View PDF](#)

化学物質のリスク評価は時間のかかるプロセスであり、すべての化学物質がタイムリーに評価され、規制されるように最適化する必要がある。この移行は、*in silico* 人工知能 (AI) /機械学習 (ML) モデルの有用なアプリケーションによって促進される可能性がある。しかし、リスク評価における AI/ML モデルの実装は遅れている。ほとんどの AI/ML モデルは、機械的な説明可能性を欠いた「ブラックボックス」と見なされており、リスク評価者が予測を十分に信頼できない原因となっている。

本論文では、AI/ML モデルの規制当局への受け入れに不可欠な要素としての「信頼」を探求する。技術的側面と技術を超えた側面を含む信頼の要素の概要を提供し、リスク評価者による信頼を構築するために最も重要と考えられる要素を強調する。結果は、リスク評価者と計算モデラーに対して、AI/ML モデルの将来の開発に関する次のような推奨事項を提供する。1) モデルをシンプルで解釈可能なものにする;2) データとデータキュレーションに透明性を提供する;3) 範囲/意図された目的を明確に定義し、伝達する;4) 採用基準を定義する;5) モデルをアクセス可能で使いやすいものにする;6) 実際の設定で付加価値を実証する;7) 学際的な設定に参加する。これらの推奨事項は、規制目的での AI/ML モデルの信頼と受け入れを促進するために、将来の開発で認識されることが理想的である。

Risk assessment of chemicals is a time-consuming process and needs to be optimized to ensure all chemicals are timely evaluated and regulated. This transition could be stimulated by valuable applications of *in silico* Artificial Intelligence (AI)/Machine Learning (ML) models. However, implementation of AI/ML models in risk assessment is lagging behind. Most AI/ML models are considered 'black boxes' that lack mechanistical explainability, causing risk assessors to have insufficient trust in their predictions.

Here, we explore 'trust' as an essential factor towards regulatory acceptance of AI/ML models. We provide an overview of the elements of trust, including technical and beyond-technical aspects, and highlight elements that are considered most important to build trust by risk assessors. The results provide recommendations for risk assessors and computational modelers for future development of AI/ML models, including: 1) Keep models simple and interpretable; 2) Offer transparency in the data and data curation; 3) Clearly define and communicate the scope/intended purpose; 4) Define adoption criteria; 5) Make models accessible and user-friendly; 6) Demonstrate the added value in practical

settings; and 7) Engage in interdisciplinary settings. These recommendations should ideally be acknowledged in future developments to stimulate trust and acceptance of AI/ML models for regulatory purposes.

## Research article

### **GARD<sup>TM</sup>skin and GARD<sup>TM</sup>potency: A proof-of-concept study investigating applicability domain for agrochemical formulations**

Marco Corvaro, Joseph Henriquez, Raja Settivari, Ulrika Mattson, ... Sean Gehen

Article 105595

皮膚感作性物質の有害性評価のためのいくつかの New Approach Methodologies (NAM) が公式に検証されている。しかし、特定の製品クラスに対するそれらの適用性に関するデータは限られている。本プロジェクトの目的は、農薬製剤の製品クラスに対する GARD<sup>TM</sup>skin および GARD<sup>TM</sup>potency の適用領域に関する初期証拠を提供することであった。

この概念実証のために、30 の液体および 12 の固体農薬製剤の有害性予測を GARD<sup>TM</sup>skin で試験した。感作性物質として予測された製剤は、GHS 皮膚感作性区分を決定するために GARD<sup>TM</sup>potency アッセイでさらに評価された。選択された製剤は、業界の製品を代表する製品タイプ、有効性グループおよび感作有害性クラスであった。

GARD<sup>TM</sup>skin の性能は、既存の in vivo 動物データと結果を比較することによって推定された。全体的な精度、感度および特異度は、それぞれ 76.2% (32/42)、85.0% (17/20) および 68.2% (15/22) であり、液体製剤の予測率は固体制剤と比較してわずかに高かった。GARD<sup>TM</sup>potency は、正しく予測された 17 種類の感作物質のうち 14 種類を正しくサブカテゴリー化した。一致の欠如は、組成分析またはボーダーライン応答分析によって正当化された。

結論として、GARD<sup>TM</sup>skin および GARD<sup>TM</sup>potency は、この最初の概念実証研究において十分な性能を示し、農薬製剤が試験方法の適用範囲内にあることを支持した。

Several New Approach Methodologies (NAMs) for hazard assessment of skin sensitizers have been formally validated. However, data regarding their applicability on certain product classes are limited. The purpose of this project was to provide initial evidence on the applicability domain of GARD<sup>TM</sup>skin and GARD<sup>TM</sup>potency for the product class of agrochemical formulations.

For this proof of concept, 30 liquid and 12 solid agrochemical formulations were tested in GARD<sup>TM</sup>skin for hazard predictions. Formulations predicted as sensitizers were further evaluated in the GARD<sup>TM</sup>potency assay to determine GHS skin sensitization category. The selected formulations were of product types, efficacy groups and sensitization hazard classes representative of the industry's products.

The performance of GARDskin was estimated by comparing results to existing *in vivo* animal data. The overall accuracy, sensitivity, and specificity were 76.2% (32/42), 85.0% (17/20), and 68.2% (15/22), respectively, with the predictivity for liquid formulations being slightly higher compared to the solid formulations. GARDpotency correctly subcategorized 14 out of the 17 correctly predicted sensitizers. Lack of concordance was justifiable by compositional or borderline response analysis.

In conclusion, GARDskin and GARDpotency showed satisfactory performance in this initial proof-of-concept study, which supports consideration of agrochemical formulations being within the applicability domain of the test methods.

## Research article

### **Evaluating scientific confidence in the concordance of *in vitro* and *in vivo* protective points of departure**

En-Hsuan Lu, Lucie C. Ford, Zunwei Chen, Sarah D. Burnett, ... Weihsueh A. Chiu

Article 105596

哺乳類の *in vivo* 実験動物試験への依存を減らすという約束を果たすために、新しいアプローチ方法 (NAM) は、規制上の意思決定のための信頼できる基礎を提供する必要がある。しかし、*in vitro* NAM ベースの出発点 (POD) を開発するための以前の試みは、例えば、U.S.EPA の ToxCast からの POD は、より保守的 (保護的) であるが、従来の *in vivo* 研究との相関性が低いように見えるなど、様々な結果をもたらした。ここでは、*in vivo* POD の不均一性を減らし、種差を考慮し、*in vitro* POD の生物学的関連性を高めることによって、この不一致に対処することを目的とした。しかし、*in vivo* POD のためのベイジアンモデル平均化に基づくベンチマーク用量モデリング、種間調整のための相対成長スケーリング、および複数の人工多能性幹細胞由来モデルとヒト関連 *in vitro* アッセイを組み合わせた場合にのみ、*in vitro* と *in vivo* の一致が改善されることを見出した。さらに、利用可能なサンプルサイズはわずか 15 化学物質であり、結果として得られた一致レベルは、相関係数 < 0.5、予測間隔が数桁に及ぶという、公平なものであった。全体として、本研究は、*in vitro* NAM に基づく POD の一致を高め、それによって科学的信頼性を高めるいくつかの方法を示唆する一方で、規制上の意思決定に使用するための予測精度と精度における課題も浮き彫りにした。

To fulfil the promise of reducing reliance on mammalian *in vivo* laboratory animal studies, new approach methods (NAMs) need to provide a confident basis for regulatory decision-making. However, previous attempts to develop *in vitro* NAMs-based points of departure (PODs) have yielded mixed results, with PODs from U.S. EPA's ToxCast, for instance, appearing more conservative (protective) but poorly correlated with traditional *in vivo* studies. Here, we aimed to address this discordance by reducing the heterogeneity of *in vivo* PODs, accounting for species differences, and enhancing the

biological relevance of *in vitro* PODs. However, we only found improved *in vitro*-to-*in vivo* concordance when combining the use of Bayesian model averaging-based benchmark dose modeling for *in vivo* PODs, allometric scaling for interspecies adjustments, and human-relevant *in vitro* assays with multiple induced pluripotent stem cell-derived models. Moreover, the available sample size was only 15 chemicals, and the resulting level of concordance was only fair, with correlation coefficients <0.5 and prediction intervals spanning several orders of magnitude. Overall, while this study suggests several ways to enhance concordance and thereby increase scientific confidence in *in vitro* NAMs-based PODs, it also highlights challenges in their predictive accuracy and precision for use in regulatory decision making.

## Research article

### **Safety assessment for nail cosmetics: Framework for the estimation of systemic exposure through the nail plate**

Xuejun J. Yin, Nicola J. Hewitt, Steffen Erler, Paul Bryson, ... Carsten Goebel

Article 105588

[View PDF](#)



c

## Correspondence

### **Dealing with substances with no defined molecular weight in non-animal assays for skin sensitization. A comment on “Plant extracts, polymers and new approach methods: Practical experience with skin sensitization assessment” ()**

David W. Roberts

Article 105582

## Correspondence

### **Response to the Letter to the Editor by David W Roberts “Dealing with substances with no defined molecular weight in non-animal assays for skin sensitization. A comment on “Plant extracts, polymers and new approach methods: Practical experience with skin sensitization**

**assessment” 0”**

Susanne N. Kolle

Article 105593

Mirait Translate AEIC trial