PrecisionTox D6.5: Report on the Capacity of Legal and Regulatory Frameworks to Accommodate NAMs

Aleksandra Čavoški, Laura Holden, Robert Lee*

With contributions from Louis Dawson, Scott Glaberman,** and Rubina Sultan-Chaudhary**

Executive Summary

The goal of PrecisionTox is to advance safety assessment of chemicals without the use of animal testing by establishing a new, 3Rs-compliant, cost-effective testing paradigm for chemical safety assessment ('Precision Toxicology'). Millions of animals are used each year in research within the EU, thousands of which are used for regulatory purposes associated with various chemical regulations and pharmaceuticals. In needing to address the risks from both the chemical backlog and substances new to the EU market, NAMs provide an opportunity to incorporate developing science. This report undertakes a critique of key risk assessment models, such as the framing of chemical testing as either protective or predictive and the focus of different jurisdictions on either hazard or exposure. The review serves to illustrate the inextricable links between these debates and suggests that value can be added from NAMs, which is not fully recognised by EU law as currently drafted, applied, or interpreted. A review of judicial interpretation of legislation through an analysis of toxicological court cases relating to the utility of chemical safety testing methods reveals significant complexities for all actors involved in the decision-making process. This review questions the extent to which the last resort principle is applied in oversight of regulatory decision making, and it exposes the wider political accountability of the decision-makers involved. Finally, this report provides an analysis of key legislative instruments in the chemical domain to uncover the extent that NAMs are recognised in EU regulatory structures. While the testing requirements lean very heavily on traditional animal testing, there does appear to be greater scope for the use of NAMs under REACH and CLP, which defies general opinion. To

*Birmingham Law School, University of Birmingham, Edgbaston, Birmingham, B15 2TT, United Kingdom, a.cavoski@bham.ac.uk; l.holden.1@bham.ac.uk; r.g.lee@bham.ac.uk
**Biosciences, University of Birmingham, Edgbaston, Birmingham, B15 2TT, United Kingdom

improve the regulatory acceptance of NAMs, we recommend clearer policy direction to address a key issue of interpretation of what determines adversity. This may be achieved to an extent by utilising soft law guidance rather than hard law legislative change, coupled with applying lessons learnt from the Cosmetics Regulation and pharmaceutical sectors' pursuit of alternative methods to animal testing.

Contents

1. Introduction

The goal of PrecisionTox is to advance safety assessment of chemicals without the use of animal testing by establishing a new, 3Rs-compliant, cost-effective testing paradigm for chemical safety assessment ('Precision Toxicology') that identifies molecular key event (KE) biomarkers predictive of chemically induced adverse health effects in humans and facilitates their uptake into regulatory and industry practice. An objective of PrecisionTox is regulatory analysis and application, which includes identifying opportunities for applying Precision Toxicology within existing regulatory structures.

It is vital to ensure that NAMs can be accommodated within the prescribed chemical risk assessment processes within the EU. To ensure this, a doctrinal study of EU legislation and case law has been carried out, based on the EUR-Lex and other legal databases. This has enabled the mapping of chemical risk assessment in EU legal frameworks to assess where and how NAMs can be utilised within present regulatory frameworks.

Additionally, a review to critique current risk assessment models in terms of regulatory efficacy has been undertaken, drawing upon academic literature, major jurisprudence of the Court of Justice of the European Union (CJEU) and the General Court, and a comparative analysis of different legal systems to highlight potential weaknesses in present regulation. The hope is that the scientific work of the PrecisionTox project can be drawn down to address such weaknesses thereby highlighting the challenges facing the present regulation. The aim is to strengthen regulatory risk governance by supporting and informing a programme of targeted withdrawal of animal testing while helping to shape a reformed regulatory framework in the light of a review of the capacity of the existing hard law structures to exploit the potential of NAMs in ensuring chemical safety. Earlier research on sociotechnical barriers to be overcome to create a climate

that is ready to accommodate and employ developed NAMs[1] suggested that legal structures did at times act as a technical barrier to their take up into regulation.

This report begins with an analysis of the present use of animal testing (Chapter Two). Chapter Three then provides a critique of a range of risk assessment models, while reviewing the regulatory efficacy of these varying approaches. Following this, Chapter Four provides a review of relevant case law with the aim of exploring how the EU courts are currently interpreting EU legislation when reviewing challenges surrounding dossier submissions. Chapter Five is an analysis of the current legal frameworks associated with industrial chemicals (REACH), classification, labelling, and packaging (CLP), biocides (BPR), plant protection products (PPP), cosmetics, food contact materials (FCM), and pharmaceuticals,[2] to identify quite how receptive such legislation is in its possible resort to NAMs. The report ends with proposed actions for the PrecisionTox Work Packages to incorporate as the project reaches its conclusion and to inform the remaining deliverables, such as that regarding the development of soft law instruments.

2. Animal Use Data

The 'Animal Use Reporting – EU System' (ALURES) database of the European Commission is the system for reporting animal use in the European Union (EU).[3] At the time of writing, the most recent data is that from 2022, for the 27 Member States of the EU and Norway. This reports the total number of animals used for research, testing, routine production, and education and

[1] Aleksandra Čavoški, Laura Holden, Robert G Lee, 'Report on Socio- Technical Barriers to the Uptake of NAMs' (*PrecisionTox,* 26 January 2024) <https://precisiontox.org/wp-content/uploads/2024/02/D6.1-Report-on-Socio-Technical-Barriers-26Jan.pdf> accessed 22 October 2024; Aleksandra Čavoški, Laura Holden, Leonie Mueller, Robert G Lee, 'Balancing Chemical Safety and Animal Welfare Considerations in the Application of New Approach Methodologies for Chemical Safety Assessment' (manuscript submitted); Aleksandra Čavoški, Laura Holden, Robert G Lee, 'The Place of Law in Technology Transitions: A Case Study from Chemical Risk Assessment' Journal of Environmental Law (manuscript accepted).

[2] Regulation (EC) No. 1907/2006 on the Registration, Evaluation, Authorisation and Restriction of Chemicals [2006] OJ L 396/1 ('REACH'); Regulation (EC) No. 1272/2008 on Classification, Labelling and Packaging of Substances and Mixtures [2008] OJ L 353/1 ('CLP'); Regulation (EU) No.528/2012 Concerning the Making available on the Market and Use of Biocidal Products [2012] OJ L 167/1 ('BPR'); Regulation (EC) No 1107/2009 Concerning the Placing of Plant Protection Products on the Market [2009] OJ L 309/1 ('PPP'); Regulation (EC) No 1223/2009 on Cosmetic Products [2009] OJ L 342/59 ('Cosmetics Regulation'); Regulation (EC) No 1935/2004 on Materials and Articles Intended to Come into Contact with Food [2004] OJ L 338/4 ('FCM'); and Directive 2001/83/EC on the Community Code Relating to Medicinal Products for Human Use [2001] OJ L 311/67 ('pharmaceuticals').

[3] Available at: <https://webgate.ec.europa.eu/envdataportal/content/alures/section1_number-of-animals.html> accessed 22 January 2025.

training to be 8,385,397 animals. The year before, in 2021, the figure was 9,406,233. Prior to this, in 2020, the figure was 7,938,064 animals. Earlier figures then include the UK as a Member State, and it could be suggested that Covid-19 may have had some effect on figures: in 2019 animal use was 10,260,822 and in 2018 it was 10,572,305. Therefore, with varying parameters and global conditions, it is difficult to draw any conclusion on trend, except to recognise that the use of animals is in the millions.

When these figures are narrowed to animal use for regulatory purposes, Table 1 presents the data from ALURES for 2020 to 2022 by type of legislation and states the species most used. These dates are used as the three most recent years, and because they each provide data for 27 EU Member States and Norway. Industrial chemicals, plant protection, and biocide legislations are reported separately and therefore appear in the table separately. Figures for FCM are reported within a wider category of food legislation, however these are still provided below for reference, and it is separate from reporting under feed legislation. Animals used for medicinal products for human use are reported separately from that for veterinary use and that for medical devices, and human medicinal products are a focal point of this report in relation to pharmaceuticals so are reported in the table.

Table 1: Uses of Animals for Specified Legislation			
Type of Legislation	**Year**	**Animals Used**	**Predominant Species**
Industrial Chemicals	2022	151,932 (23.8% increase)	Rats 64.8%
	2021	165,086	Rats 74.7%
	2020	122,736	Rats 73.0%
Plant Protection Products (PPP)	2022	35,895 (46.6% decrease)	Rats 41.3%
	2021	55,798	Rats 36.2%
	2020	67,174	Rats 35.0%
Biocides (BPR)	2022	4,069 (8.4% decrease)	Zebrafish 34.6%
	2021	2,578	Mice 34.1%
	2020	4,442	Rats 45.0%
FCM (Food, including FCM)	2022	8,141 (71.3% decrease)	Mice 64.5%
	2021	16,079	Mice 87.9%
	2020	28,326	Mice 94.0%
Human Medicinal Products	2022	515,344 (32.1% decrease)	Mice 53.0%
	2021	673,409	Mice 58.1%
	2020	758,902	Mice 60.9%

Animal use under industrial chemicals legislation increased by 23.8% between 2020 and 2022 and the large majority of this was using rats. While rat use is declining, rabbit and mice use have seen small percentage increases (9-12% and 2-3% respectively).

Animal use for PPP has declined by nearly half between 2020 and 2022. Although animal use is still in the tens of thousands, the percentage share of use of rats has increased, while other species type has fluctuated: for example, the use of Zebrafish was at 10.5% in 2020; 26.1% in 2021; and reduced to 5.88% in 2022.

For biocides, a sector with far fewer (albeit still thousands) of animals used, there was an 8.4% decrease in animals used from 2020 to 2022. The predominant species has become that of the alternative model species zebrafish, with rat use decreasing from 45% in 2020 to 13.5% in 2022. Xenopus has also seen an increase in use, from no significant percentage in 2020 to 29.6% in 2022.

While mice remain the major species of use for FCM, animal use overall has greatly decreased by 71.3% from 2020 to 2022, a trend that stretches across the three years observed, from 94.0% in 2020 down to 64.5% in 2022. Unfortunately, reliance on the use of rats appears to have largely balanced out the reduction in mice by, increasing from 5.4% in 2020 to 33.0% in 2022.

Finally, human medicinal products account for the greatest use of animals within the legislation under review here but also show a decrease in animal use by 32.1% from 2020 to 2022. However, animal use remains largely mammal based with Zebrafish use only making it above a percentage point use in 2022 (1.7%).

Varying levels of animal use can be observed between the legislative sectors, with most sectors showing decreasing use of animals, with the steepest decrease being that for FCM. Yet, even where animal use is decreasing, the use of mammals as a share of animal use is increasing, as is the case under PPP, FCM, and human medicinal product legislation. Animal use for industrial chemicals has increased and the use of mammals is consistent. The most optimistic data was under the biocide legislation, where animal use and the percentage share of mammals within that figure both decreased.

These statements are based, however, on a review of a limited dataset of three years due to variations in the base data (due to, for example, the UK leaving the EU), and conclusions are limited as to their cause because of the many external factors, such as the impacts of the Covid

pandemic on the relevant sectors. However, it is possible to suggest that reductions in animal use are due to a greater use of alternative methods under the umbrella term of 'NAMs'.

In 2020, ECHA received 14,928 registration dossiers (including updates) under REACH;[4] in 2021 ECHA received 14,928 registration dossiers (including updates);[5] and in 2022, ECHA received 13,530 registrations (including updates).[6] As these registration figures remain consistent, this indicates that efforts to reduce animal testing have yet to bear fruit, as animal use data has increased by 23.8% from 2020 to 2022, but registration dossiers have seen a 9.4% decrease in the number of submissions to ECHA over the same period.

We were only able to access anecdotal suggestions as to the changes in animal use, owing to factors such as the cyclical renewal process of BPR and PPP. We suggest this as an area of further research, to help determine the true causes as to the varying figures and provide greater depth to the understanding of headline animal use numbers. It is also apparent that the total figures are only indicative, as Busquet et al. note in stating that foetal forms of mammals (pups) are excluded from reporting requirements.[7] The authors state that 'far more pups than adult animals' are used in two-generation studies, claiming up to thousands of animals from 20 each of male and female adults are under-reported. A paper by Rovida et al. suggests under-reporting due to testing conducted outside the EU.[8] However, what the authors do note, which is pertinent for meaningful reduction in animal use, is that the greatest number of animals are used for repeated dose toxicity, reproductive toxicity, and development toxicity, alongside ecotoxicological tests requiring fish.

Publicly available data on animal use in chemical risk assessment therefore provides only a partial picture. However, animal testing for regulatory purposes remains in the thousands, and the greater deployment of NAMs ought to be able to reduce this figure. The next chapter

[4] ECHA 'Annual Report 2020' (*European Chemicals Agency,* April 2021) available at: <https://echa.europa.eu/documents/10162/7362407/annual_report_2020_en.pdf/09d078c5-ff40-6737-3e4c-41dea91a7738?t=1619715877119> accessed 22 January 2025, p56.

[5] ECHA 'Annual Report 2021' (*European Chemicals Agency,* May 2022) available at: <https://echa.europa.eu/documents/10162/11872732/mb_05_2022_2_annual_report_2021_mb65_en.pdf/7688a1e9-5d23-59fb-213c-2bd940c052ff?t=1651662515417> accessed 22 January 2025, p75.

[6] ECHA 'Annual Report 2022' (*European Chemicals Agency,* April 2023) available at <https://echa.europa.eu/documents/10162/21371921/mb_03_2023_annual_report_2022_en.pdf/d6a5b3dd-e4cc-7c99-2962-75a168b8de4b?t=1682340571095> accessed 22 January 2025, pp64, 19.

[7] Francois Busquet, and others, 'New European Union Statistics on Laboratory Animal Use - What Really Counts!' (2020) 37(2) ALTEX 167.

[8] Costanza Rovida, and others, 'View of REACH Out-numbered! The Future of REACH and Animal Numbers' (2023) 40(4) ALTEX 367.

considers the risk assessment models of the existing and NAM-based paradigms, to characterise these debates and their anticipated efficacy in reducing animal use.

3. Critique of Risk Assessment Models

With contributions from Scott Glaberman

Having considered the state-of-play of the implementation of the 3Rs based on an analysis of animal use data, this chapter reviews current risk assessment models to analyse the efficacy of the current animal-based regulatory paradigm in light of the opportunities afforded by NAMs. This critique is conducted through a series of juxtaposed debates namely:

- Hazard vs Exposure
- Prediction vs Protection
- Observing Adversity vs Developing knowledge of Modes of Action
- Assessment of Single Substances vs Groups of Substances
- Single Method Test vs NAM Test Batteries

These choices have been selected due to their recurrence throughout the empirical study of D6.1 into the barriers to the uptake of NAMs.[9] Additionally, they form key areas of challenge in scholarship, for example the framing of animal tests as being predictive has been contested by highly regarded scientists, such as the authors in Browne et al., who note that, for example, a percentage change in rat body weight is not directly applicable to humans but may be indicative of a harmful effect requiring uncertainty factors of multiplication to be applied to derive doses for humans.[10] Others note that the predictive capacity of animal tests is somewhat assumed,[11] and that where clinical data is available, NAMs have demonstrated greater predictivity.[12] Likewise, the development of adverse outcome pathways around which modes of action and integrated testing approaches may be based are a focus of JRC research and OECD papers,[13] providing a

[9] Čavoški (n1) 'Report on Socio- Technical Barriers to the Uptake of NAMs.'
[10] Patience Browne, and others, 'Adverse Effects in Traditional and Alternative Toxicity Tests' (2024) 148 Regulatory Toxicology and Pharmacology 105579.
[11] Anna J. van der Zalm, and others, 'A Framework for Establishing Scientific Confidence in New Approach Methodologies' (2022) 96 Archives of Toxicology 2865.
[12] Ibid.
[13] A. Carusi, and others, 'Adverse Outcome Pathway – Study Report' (*European Union*, 2022) available at <https://data.europa.eu/doi/10.2760/476003> accessed 22 January 2025; Daniel L. Villeneuve, and others, 'Adverse Outcome Pathway (AOP) Development I: Strategies and Principles' (2014) 142(2) Toxicological Sciences 312; OECD 'OECD Series on Adverse Outcome Pathways' available at: <https://www.oecd-ilibrary.org/environment/oecd-series-on-adverse-outcome-pathways_2415170x> accessed 22 January 2025.

structure for NAM data to be presented in place of observational endpoints. This chapter will identify areas in which NAMs have potential to strengthen regulatory risk governance towards meeting their aims through reviewing the debates derived from scholarship and interview responses before considering the role of PrecisionTox in the context of wider policy considerations.[14]

3.1 HAZARD VS EXPOSURE

Risk assessments are commonly depicted as a series of stages which include: hazard identification (i.e., is this something that has the potential to cause a harm); exposure assessment (i.e., are there scenarios where a person / the environment would come into contact with the hazard?); and risk characterisation (i.e., would the exposure in those scenarios lead to a sufficient dose that would be harmful?). However, in relation to the many thousands of chemicals circulating on a global market, jurisdictions have taken varying approaches to in the operation of the risk assessment framework. In the EU this follows the steps described above, where hazard is first identified and communicated. Therefore, notwithstanding any initially known uses of a substance, should those uses change, the fundamental knowledge of the hazard remains applicable. This perhaps reflects the fact that in the EU the onus is on industry to provide information on their substances down the supply chain using safety data sheets, so that informed decisions on the marketing of substances are based on industry data. The CLP further entrenches this, by requiring industry to classify the hazards of their substances and to label them accordingly. Conversely, in the USA, regulators need to justify requests for information about substances and even conduct testing themselves, thereby meeting the need to assuage concern of a harmful exposure.

A typical description of the EU regulatory framework for the industrial chemicals is that it represents a hazard-based system. This is not strictly true, as the hazardous endpoints to be tested for under REACH are dependent on tonnage, which is employed as a proxy for exposure. These endpoints are cumulative as tonnage increases. However, perceptions of a hazard-

[14] The empirical study involved 32 semi-structured interviews using Zoom software, conducted from January 2023 to May 2023 with participants from three groups of stakeholders involved in the risk assessment and management of chemicals, namely industry, regulators, and policy makers, across several jurisdictions. References to interviewee stakeholder groups are I=industry, R=Regulatory, P=Policy respondent. Each respondent is assigned a numerical figure within their stakeholder group, for example P1 references policy respondent 1.

focused system are reinforced through features of the regulation such as that: acceptance of exposure-based waivers are said to be hard to attain through REACH;[15] standard information requirements are based on lists of adverse endpoints; and REACH cross-references the hazard-based requirements of the CLP Regulation. Whether registrants are meeting the requirements of REACH in terms of, for example, the exposure assessments and risk characterisation required for chemical safety reports,[16] appears under-evaluated as compared to hazard testing, despite the authority of ECHA extending to ensuring that such assessments, reports, and proposed risk management measures are adequate.[17] Issues with determinations around hazards tend to be the focus of appeals to decisions, as discussed in Chapter 4 below. Following evaluation, subsequent authorisation or restriction decisions are typically by reference to the uses of a substance, yet such determinations are made on significantly fewer substances than have been registered due to resourcing constraints. Therefore, most registrants are aware that their substances will be subject to a dossier compliance check, which takes a greater hazard focus as the evaluation checks against the endpoint-based standard information requirements.

Some consider that in relation to environmental health, the use of tonnage is an acceptable tool, but it is called into question for human health effects, as humans may be exposed to small amounts of a particularly hazardous substance for which the full harmful effects remain uncertain. For example, reproductive toxicity is not a standard information requirement for substances manufactured or imported in quantity of one or more tonnes per year, and there are no requirements for substances under this tonnage. This does not mean that industry is not self-informed as to the hazards of their substances or is neglecting appropriate risk management measures in their suggested use and handling, but merely that the protective aims of legislation are not guaranteed. Given the convoluted nature of industrial chemical supply chains, it is generally considered (although not without contention) to be too great a task to gather actual exposure data.

The impact regulations have on each other can also support the claim that the EU system is hazard-based. Under REACH, reproductive toxicity information is not required until the Annex VIII 10 tonne threshold has been reached,[18] yet CLP classifications (which includes reproductive

[15] Interviewees I2, I3, I7, I8, I9, P2, P4, P6.
[16] REACH Article 14 and Annex I.
[17] REACH Article 41.1(c).
[18] REACH Annex VIII, point 8.7.

toxicity)[19] apply regardless of tonnage.[20] A REACH restriction states any substances identified as CMRs (carcinogens, mutagens, and reproductive toxicants) shall not be placed on the market or used for supply to the general public.[21] Outside of this public use restriction, identified CMRs are still considered to be substances of very high concern, which is a justification for them being the subject of REACH authorisation procedures.[22] Yet while authorisations in general are a risk management measure to regulate use and exposure, the inclusion of a substance on the candidate list is met with concern by industry that their use may be limited, which then acts as a deterrent based on a hazard identification. [23] Additionally, while the CLP Regulation does not apply to cosmetics, under the Cosmetics Regulation CMRs are also banned from cosmetics, regardless of exposure. From a regulatory perspective however, these measures could be said to demonstrate the effective application of legislation; it is not that industry intends to harm health, but because action is taken on particular hazards this is a harmonised approach to ensuring that exposure cannot be harmful, and therefore such hazard-based measures are a protective mechanism for the EU.

To further defend the consideration of exposure in the EU system, exposure scenarios are required to varying degrees across the chemical safety legislation. This is apparent, for example, in the Chemical Safety Reports of substances over 10 tonnes per year in REACH and for worker, operator, and resident exposure estimates for plant protection products and their active substances. Overall, though, while sector-specific legislation typically involves an exposure-based approach, it is the hazard focus of the horizontally applied CLP legislation that drives the EU's focus on hazard. [24]

Through the empirical study conducted in D6.1 on the barriers to the uptake of NAMs, interviews uncovered a range of views on these varying approaches to risk assessment. Interviews included global industries, as well as non-EU regulators from jurisdictions taking a more exposure-based approach. Here it is observed that, to manage regulatory workloads, chemical safety frameworks typically implement two systems, one addressing 'existing' chemicals in their jurisdiction, and those 'new' substances being brought to market. A key utility of NAMs in exposure-based

[19] CLP Annex 1, section 3.7.
[20] ECHA 'Introductory Guidance on the CLP Regulation' (*European Chemicals Agency,* January 2019) available at <https://echa.europa.eu/documents/10162/2324906/clp_introductory_en.pdf/b65a97b4-8ef7-4599-b122-7575f6956027?t=1547546145023> accessed 22 January 2025, p31.
[21] REACH restriction, Annex XVII, item 28.
[22] REACH Article 57.
[23] Kristina Nordlander, Carl-Michael Simon, Hazel Pearson, 'Hazard v Risk in EU Chemicals Regulation' (2010) 1(3) European Journal of Risk Regulation 239.
[24] ECHA (n20) p25.

systems is their application in modelling exposure. While NAMs can also provide hazard information, the likelihood that this will be at the cellular level or smaller draws criticism, in that cells do not reflect organs, systems, or whole organisms.

> **NAMs give more upstream information of molecular initiating events or preceding events leading to the toxic effect, yet the [CLP] is a hazard-based system, which is based on hazards, which is usually not the information you've directly derived from NAMs, which is a problem. Now, if you have like in the US a more risk-based system, then it's not that much of a problem because you just have a point of departure. There are still enough problems with dosing, but there's no fundamental problem like you have with hazard legislation.**[25]

Rather than viewing NAMs mainly as a replacement of animal models (as is done with hazard identification, being the first step in hazard-based systems), exposure-based systems can utilise NAMs in areas that animal testing may struggle to answer, and therefore exhibit a greater capacity to employ NAMs as data-rich tools when the risk assessment involves exposure, rather than having to question how NAMs might help replace animal tests. In so doing, exposure-based approaches can therefore reduce the need for animal testing, as this is not necessary where exposure is demonstrated to be sufficiently low, so that 3R principles are better engaged.

At this point, the debate returns to the information held on the uses of substances. Where these are not fully declared or updated, hazard proponents could argue that potential harm could then be caused. When considering the level of trust inherent in the REACH processes, which places the onus on industry to communicate risk and where regulatory checks through substance evaluations are so onerous that levels of this scrutiny are low (leaving many thousands of substances currently under-evaluated), the risk of examining exposure as an early step in the risk assessment does not seem a large departure from the status quo. In the absence of the separation of new and existing substances in the EU system, information on exposure may provide an important means of prioritisation. In contrast, other than some rudimentary ranking, industry is primarily responsible for providing data on which regulatory decisions must rest, which might constitute an inherent risk within the present system.

This leads to questions surrounding the definition of NAMs. As the science of extrapolations even from animal tests is basic (albeit continually refined), the comprehension of exactly what is needed to be able to truly understand health effects might involve more than just introducing a

[25] Interviewee I11.

chemical to an organism in a test. Rather than simply considering the function of NAMs as a replacement for animal testing or even re-opening debates whether hazard or exposure should be considered first, the risk assessment model may be more transparently served by a more radical shift to an improved system. Such a system might reflect the increasing understanding of the potential adverse outcomes of exposure to substances, utilising modern science and technology to better appreciate exposure and toxicokinetics. It might also better identify hazards using non-animal methods such as the high-throughput and computing capabilities of bioinformatics, which may hold greater potential in revealing hitherto unknown or underexplored issues.

Even in jurisdictions that operate exposure-led scrutiny, the data on which this is based will be any available hazard information such as that gleaned from international hazard classification of substances. This raises the question of whether in the instance a truly new chemical under assessment for which no hazard information was available, could a regulator be confident to approve a use that is said not to give rise to any exposure? The hazard-based approach does at least nod to a chance of a 'what if' scenario and therefore it can be argued is ultimately more protective of its citizens, allowing for CLP labelling such as prevention precautionary statements (for example, 'wash contaminated clothing before reuse') and emergency procedures under REACH.

Interviewees during the early empirical research on barriers to the uptake of NAMs felt that exposure is under-used in the industrial chemical legislation and by the ECHA regulators (I2, I3, I8, I9, P6),[26] whereas it is thought to be considered more widely employed in the governance of pesticides and chemicals in food by EFSA regulators, and in cosmetics regulation (R1, P4). One respondent felt that considering exposure 'unlocks a whole world of tools that you can use' (I3) and policymakers acknowledged this may reduce animal testing (P2). One explanation for the limited consideration of exposure under REACH was attributed as resulting from a focus on worker protection. Another frequent argument was that given the wide range of industrial or consumer uses to which substances can be put, regulators can feel greater confidence in a regulatory framework with a focus on intrinsic hazardous properties (I6, I7, P3, P4). This is compounded by the focus of CLP fixing on hazard (P3). The CLP is seen as the 'underlying' regulation rather than merely a means of hazard classification as a hazard becomes apparent, despite the formal adoption of the United Nations' 'Globally Harmonized System' (GHS) of classification and labelling of chemicals coming two years after the passage of REACH. In spite

[26] Čavoški (n1) 'Report on Socio- Technical Barriers to the Uptake of NAMs'.

of the status of GHS as a global framework, it is noteworthy that jurisdictions apply GHS variably: for example, the USA applies GHS via occupational safety through a hazard communication standard, but it does not apply to pesticides, nor to some consumer goods.[27] The status of CLP in the EU may be seen as a case of the 'tail wagging the dog'. However, the preamble of CLP notes the intention of the EU to be at the 'forefront' of countries adopting GHS into legislation,[28] and as with the CLP's forerunner,[29] the focus on hazard appears to be a deliberate choice by the EU. It is therefore unlikely that this stance will change a great deal, and the EU remains intent on taking the lead on hazard identification.

Such a focus on hazard does have regulatory advantage, in that appropriate risk management measures can be applied downstream according to use as this becomes apparent. Therefore, the first step of hazard identification has been addressed in a standardised way (through CLP), with sector-specific approaches to exposure and risk then dealt with individually. From a regulatory viewpoint, knowledge of hazard allows users to consider the chemical safety reports and safety data sheets to inform controlled use, with formal regulatory action only required thereafter for higher risk substances in the case of application subject to authorisation or restriction. In terms of PrecisionTox, NAMs may contribute to this approach by providing hazard information in the form of data concerning bioactivity and via the potential for high throughput. It was noted by interviewees that the difficulty in translating exposure of model organisms (and models in general) to human exposure (I4, R4, R7, R9, R10, P1) and anchoring to harmful effects will be crucial for building confidence in the relevance of biomarker readouts for regulatory decision making.

Although the hazard versus exposure debate is likely to endure, it seems unlikely there will be a wholesale shift to exposure in the future, with the centrality of CLP likely to preserve the hazard focus. The hazard-based approach of the EU is already different in nature to that of exposure-led assessment found in certain other jurisdictions, such as the USA, Canada, and Australia. Yet, the European single market is sufficiently large that this difference has not critically hampered the chemical industry thus far, again entrenching the existing approach.

[27] ChemSafety Pro, 'GHS Implementation in USA' (December 2015) available at <https://www.chemsafetypro.com/Topics/USA/GHS_in_USA_SDS_label.html> accessed 22 January 2025.
[28] CLP Preamble note 7.
[29] Directive 1999/45/EC Concerning the Approximation of the Laws, Regulations and Administrative Provisions of the Member States Relating to the Classification, Packaging and Labelling of Dangerous Preparations [1999] OJ L 200/1.

3.2 PREDICTION VS PROTECTION

Often raised in debate about chemical regulation, and by extension NAMs, is the question of whether the aim is to be predictive or protective. While the legislation seeks to be protective of human health and the environment, to an extent, the legislation pursues both objectives. Being unable to directly test the effects of substances on humans, testing reverted to animals as a proxy for humans. A recent paper by Browne, et al., suggests the possibility of rethinking how we frame the results:

> **It is helpful to reflect on how *in vivo* data are generally used in regulatory risk assessments. While rodent models are assumed to be relevant and predictive of effect in humans, this is only to a degree. For example, a 25 % decrease in the body weight of a rat is not assumed to predict a 25 % decrease in a human exposed to an extrapolated dose of the same chemicals. Safety assessment (or uncertainty) factors are used to derive non-cancer reference doses for humans below which an adverse effect is unlikely to occur. The specific safety assessment factor applied to observed POD varies across groups and exposure scenarios, and are designed to account for uncertainty in species extrapolation, population variability, life stages, and exposure durations. In fact, mammalian toxicity data are ideologically considered to be predictive of human responses, but practically used in protective frameworks. Rather, acknowledging that the decades of accumulated animal reference data may be considered protective of adverse chemical effects rather than predictive of specific chemical effects permits development of new methods and approaches that do the same…Retrospective analyses indicate that animal studies are not highly predictive of organ-specific effects or adversity in humans; however, animal test data have a strong negative predictive value for of effects in humans. Therefore, NAMs benchmarked against *in vivo* animal data, even when the animal reference data are associated with a specific type of adversity, may be more effectively interpreted as protective, as opposed to predictive.**[30]

Therefore, in the absence of human testing, examining the outcomes of NAMs in relation to the data we do have (from historic animal testing), continues to serve a protective purpose. This then leads to policy questions of how exact such results must be. For example, how many orders of

[30] Browne (n10).

magnitude difference could still be considered protective to be an acceptable NAM, considering the data against which the NAM data are being assessed is not truly predictive itself?

An aim of many NAM developers is to increase the human relevance of a method. Understanding of shared toxicity pathways and biology across other species serves not only to reduce the criticism of simplicity that is directed at single-cell NAMs, but such an approach can also indicate the boundaries of responses from multiple whole-organism model species, which helps build evidence towards relevance for humans.

A limitation of both NAMs and animal testing is the relevance of predictions at the population level. Organisms are in constant flux, with different genes being expressed in response to changing conditions in their environments. Therefore, any test system is open to questions about how predictive it can be across many individuals in a population and how results can be scaled across such a range when lab-based organisms exhibit different tolerances, may be closely related, with all populations holding a spectrum of sensitivity. This is perhaps an area to which big data and AI can contribute.

Here we may combine arguments for an exposure-focused risk assessment with that of the notional split of assessing risks for humans and for the environment. Proponents for exposure-based risk assessments maintain that the level of exposure at which harmful levels occur are not typically levels to which humans will be exposed. Therefore, the parameter needing prediction is the exposure level. Beyond that level, if bioactivity is measured with a clear linkage to human relevance and actual exposure, then a risk assessment will be required. In the absence of human relevance, effects on the environment alone may be considered.

Further limitations are also apparent in the ability to conduct meaningful tests for the assessment of mixtures and particularly of realistic mixtures found in the environment. Monitoring and testing are both hindered by the variable levels of contaminants likely at the same sampling point due to temporal and environmental variables, such as seasonality and rainfall. In addition, there is little environmental or human biomonitoring undertaken in the EU, leading to crude assessments of safety levels and condition. Nevertheless, limiting exposures to anywhere near effect levels offers protective, even if not predictive, cover. Policy decisions can again apply based on the probabilities considered acceptable, not dissimilar to a cost benefit equation such as that for national healthcare decisions on the pharmaceutical provision.

With many limitations highlighted in the search for an effective form of assessing risk, this results in an imperfect system, whether the intention is to be predictive or protective. Some respondents

were concerned about the risk of losing substances due to overregulation, in attempts to protect rather than predict. For example, if in future the identification of any bioactivity was to become associated with taking formal risk management action, despite that bioactivity not yet having been anchored to a specific harmful effect (I1, I12, R2, R5). This is also an important point for the development of the PrecisionTox approach: identifying the point of departure was seen as important by many respondents. Yet questions of how predictions of human harm could be made were raised, as it was felt by some interviewees that the animal model is a poor predictor of human safety and therefore not a suitable benchmark for NAMs (I2, I6, R5), even if its combination with multiplications for uncertainty factors was acknowledged as being protective (I2). The application of AOPs was highlighted by some as representing NAM relevance (R3, P3), and predictive models were thought to have greater potential for addressing environmental endpoints (I4), which may be a 'path of least resistance' for early NAM acceptance.

This debate is therefore a key area for the progression of NAMs. In not being able to test chemicals on humans, the development of AOPs to illustrate how adversity occurs and to anchor the relevance of NAMs becomes influential in the replacement of the assumed (yet doubted) predictive ability of animal models. Here, the need for a strong science surpasses the ethical driver for using NAMs. At the same time, clear policy signals are required to determine when enough information is 'enough' for the relevant level of protection.

3.3 OBSERVING ADVERSITY VS DEVELOPING KNOWLEDGE OF MODES OF ACTION

As an elaboration of a point touched on above in the 'Hazard vs Exposure' section regarding NAMs and molecular initiating events, a criticism of the forms of data derived from NAMs is that it does not meet the legislative requirements. In some instances, the wording of legislation seems unequivocal in requiring animal testing. As an example, the data requirement for the active substances of plant protection products in section 5.3.2 of Part A of the Annex is for an 'Oral 90-day study' and states:

'The short-term oral toxicity of the active substance to rodents (90-day), usually the rat, a different rodent species shall be justified, and non rodents (90-day toxicity study in dogs), shall always be reported.'[31]

In other cases, such as REACH's requirements in the legislation for toxicological information, in Annex VIII, point 8.5 simply states 'Acute Toxicity.' While 8.5.2 refines this to 'By inhalation', it does not state specifically what is expected to be observed. For this we must turn to the guidance.

The legislative mapping for this deliverable provides a more comprehensive picture across the regulatory landscape, but ultimately it is in soft law where the apical endpoints to be observed are described. The expectation of regulators to review certain test outcomes appears to drive the interpretation of the legislation, rather than the interpretation being amenable to NAMs. This is likely because animal testing was the leading form of toxicological investigation throughout the early regulation of chemicals. However, in light of 3R principles, the legislation seems confused at best,[32] with exhortations to employ animal testing as a last resort sitting alongside express demands to test on animals. The legislation does not explicitly exclude the use of NAMs, however, and in some cases, as in REACH Annex XI, possible adaptations are described.

How well either of the approaches (observing adversity versus modes of action) actually contribute to assessing risk is open to debate. While apical outcomes observed in animals are seen as a poor proxy for humans, modes of action continue to be revealed. Modes of action are also connected with the Prediction vs Protection debate above, in terms of how much information is truly needed to be able to make protective regulatory decisions. Indeed, the focus on AOPs rather than animals leads to fears that 'something' may be missed. Having discussed above that organisms are not static, and that susceptibility varies, claiming all linkages or networks are known in an AOP may lead to the unintended consequence of a false sense of security. This in turn may limit future research on non-specific endpoints (perhaps unknown key events). As an example, the range of symptoms experienced from Covid-19, particularly extreme when accompanied by underlying health conditions, in some cases leading to long-Covid, indicates the variability of human effects.

[31] Regulation (EC) 283/2013 Setting Out the Data Requirements for Active Substances, in Accordance with Regulation (EC) No 1107/2009 Concerning the Placing of Plant Protection Products on the Market [2013] OJ L 093/1.

[32] Reference of 3Rs in Directive 2010/63 and its precursors.

Interviewees spoke of concerns about the predictive nature of knowledge of mode of action and whether biological coverage would be sufficiently extensive if attempting to replace testing on whole mammals (I1, R2, R4). In terms of the alternative model species used in PrecisionTox, the conservation of modes of action across species was questioned, with concerns of whether there may be downstream divergence (I1, R4, R5). While logical, it was suggested that a move to regulating on the basis of modes of action would be considered too great a conceptual leap for others to accept (I11, R5).

While regulatory decisions that have been based on animal testing appear to have kept us safe over the years, this is only an appearance, and we do not know and may not necessarily uncover the truth. But it is still apposite to encourage, in the light of scientific advancement, a transition to methods of greater human-relevance that might generate increased regulatory efficacy, including the meeting of ethical objectives to have recourse to the 3Rs. This does not necessarily require a retrospective analysis of decisions in a misguided attempt to apportion blame, for what at the time were justifiable decisions. However, resort to approaches such as AOPs might serve as a basis for structuring knowledge and as an indicator of pathways to harm, without which the acceptance of NAMs may be hampered further. Our legislative mapping in this deliverable may go some way to elucidating where interpretation is amenable to NAMs, which could then be strengthened in application with appropriate soft law guidance. There will need to be a willingness to accept NAM data, based on the acknowledgement that both NAMs and animal testing are imperfect systems from which hypotheses on human and population effects can never be truly tested. That knowledge of modes of action and AOPs are developing also connects this theme to the prediction vs protection debate, whereby policymakers will need to clearly signal acceptable scientific evidence.

3.4 ASSESSMENT OF SINGLE SUBSTANCES VS GROUPING SUBSTANCES FOR READ ACROSS

The use of grouping approaches is increasingly exploited, aided by the uncovering of modes of action. Typically, grouping is used by registrants as a means to avoid animal testing for the registration dossier of a substance, and informally by regulators to screen and prioritise chemicals for evaluation. Yet, there has been little regulation in reference to groups of substances. A high-profile example is that of PFAS substances, which persist in the environment. While PFOS substances within the PFAS group are restricted under the Persistent Organic

Pollutants Regulation (POPs),[33] restriction of a PFAS group under REACH, that might have a wider range of uses, is still currently being evaluated.[34]

Questions surrounding the grouping of substances in science include: the limits to the group boundaries (for example, what defines general similarity for a group and at what point does a difference require the creation of a new group); the extent to which groups should remain static or be dynamic; and the aim of grouping for positive and negative purposes (grouping undertaken by industry is generally perceived to be utilised to illustrate the absence of a harm, whereas regulatory use is to identify likely harm). Grouping also has utility limits in risk assessment, as it provides a qualitative demonstration of mechanism or behaviour, rather than a dose-response or exposure consideration.

From the empirical data, some respondents held a positive view of grouping. Grouping was seen as successful in non-EU jurisdictions, and by industry and EU policymakers, for prioritisation efficiency and avoiding animal testing (I8, I12, R4, R7, R9, P4). Others spoke of less accepting of grouping, in part because the bar of similarity was perceived as too high, due to concerns of uncertainty, leading to over-regulation (I4, R2). The advantages of grouping were appreciated more by industry which employed it for internal screening purposes (I4, I8, I12, I10).

While the utility of grouping seems widely advocated, its acceptance within the EU regulatory system is more cautious. In response to this, PrecisionTox have developed the 'Group First, Regulate Better' (GFRB) concept, to protectively formalise substance grouping based on signals of adversity (Figure 1).

[33] ECHA, 'Per- and Polyfluoroalkyl Substances (PFAS)' (*European Chemicals Agency*) available at <https://echa.europa.eu/hot-topics/perfluoroalkyl-chemicals-pfas> accessed 22 January 2025.
[34] ECHA, 'ECHA Publishes PFAS Restriction Proposal' (*European Chemicals Agency,* February 2023) available at <https://echa.europa.eu/-/echa-publishes-pfas-restriction-proposal> accessed 22 January 2025.

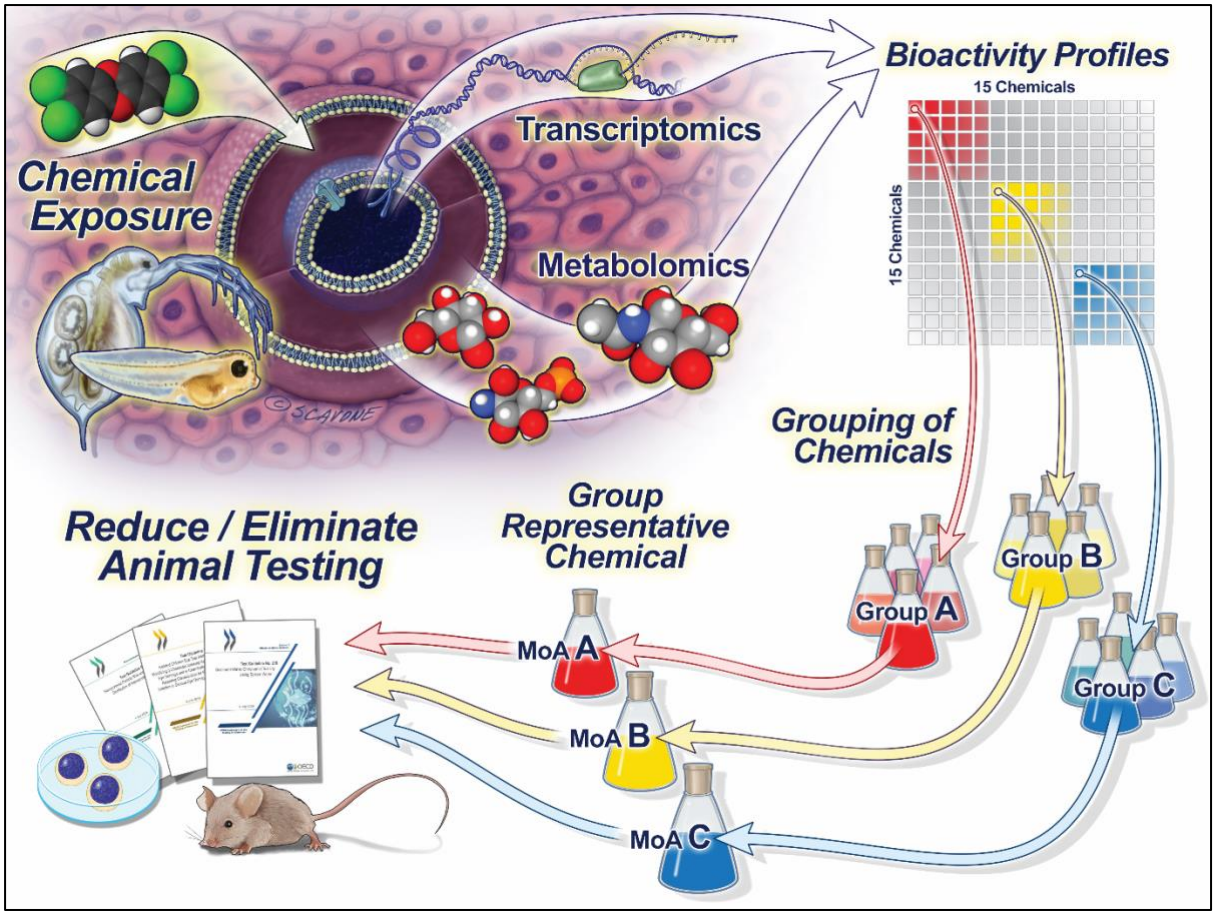

Figure 1: Representation of the 'Group First, Regulate Better' approach

GFRB utilises appropriate, 3Rs-compliant, biological test systems, which are subjected to a chemical exposure. High volumes of informative data reflecting a wide range of bioactivity enables grouping on observed biological effects, including how a chemical is metabolised, and offers an empirical dimension for chemical grouping. Group-representative chemicals are selected from groups for higher-tier toxicity testing, where bioactivity profiles are without mechanistic anchoring. The results are applied to all members of the group from which they were taken, thereby reducing the use of animal testing for all substances within the group. These groupings are further refined inductively, as data accumulates and measurement advances, which ultimately leads to an elimination of animal testing as mechanistic understanding develops. A later deliverable of PrecisionTox (D6.7) will provide a case study, applying the GFRB approach.

3.5 SINGLE METHOD TEST VS NAM TEST BATTERIES

This section refers to NAM test batteries by which is meant, inter alia, Defined Approaches (DA), Integrated Approaches to Testing and Assessment (IATA), Next-Generation Risk Assessment (NGRA), and Weight of Evidence (WoE).

The current testing paradigm typically utilises one test comprising a number of animals, from which it garners information ranging from the number of offspring to behavioural observations, and organ weight measurements. In contrast, it is generally held that one in vitro assay is unlikely to replace such a test, and instead many individual NAMs are selected and combined into sequential workflows or tiered approaches.[35] Together, these may operationalise the structuring of AOPs or span the entire assessment procedure by testing an exposure hypothesis.

While moving from ethically problematic animal testing is a noble driver, as test batteries each involve a combination of NAMs they are subject to scientific judgment as to the appropriate NAMs to use and the characteristics of a harm that need to be represented. Their value, however, is to comprehensively organise information, with an aim of not just replacing animal studies but improving shortfalls and information gaps of the existing system, as in the case of the development of the in vitro battery for developmental neurotoxicity.[36]

Respondents in our empirical study also argued that one-to-one replacements were unrealistic and that replacing the more complex endpoints could not be addressed in such a way (R3, P6). Test batteries were supported as the means to provide data from NAMs (I6, I10, R3, R7, P3, P5), with some respondents stating IATA (I6) and WoE (I2) had been submitted in dossiers. One respondent reported that the wider uptake of such batteries was limited by regulator expertise, considering the range of science and technology utilised across multiple NAM-types (I1, I10), and that regulators may continue to fall back on asking for similar representations of apical outcomes, as observed in animal studies, unless there is legislative change (I1). Some respondents were concerned about accepting NAM data and instead considered WoE-type information only as complimentary information rather than a suitable replacement for animal studies (I10, R8). Others noted the inconsistency of WoE approaches led to their being under-

[35] Helen Prior, and others, 'Reflections on the Progress Towards Non-animal Methods for Acute Toxicity Testing of Chemicals' (2019) 102 Regulatory Toxicology and Pharmacology 30.

[36] Lena Smirnova, and others, 'Revolutionizing Developmental Neurotoxicity Testing – A Journey from Animal Models to Advanced In Vitro Systems' (2024) 41(2) ALTEX 152; Magdalini Sachana and Timothy J Shafer and Andrea Terron, 'Toward a Better Testing Paradigm for Developmental Neurotoxicity: OECD Efforts and Regulatory Considerations' (2021) 10 Biology 86.

accepted by regulators and therefore under-submitted by industry (P4). As mentioned above, the legislative mapping in this deliverable may provide interpretation that can help form the basis of NAM acceptance.

It was noted that writing a Test Guideline for every context of use is not feasible (I10), not just due to the range of NAMs and developing approaches that can be combined, but also due to reports of the time-consuming nature of the task in having taken over ten years to achieve acceptance of one DA (I11, R7). Even writing guidance on how to undertake WoE has produced only 'common denominator principles,' which individuals interpret differently (P1). Since the regulatory onus is on industry to provide data, such expectations of formal case-by-case guidance are not likely to be in the spirit of the regulation as this adds burden to regulators, and there is no legal requirement for WoE to be validated (a point held as a positive by one regulator (R4)). Instead, the formalised constructs of test guideline and defined approaches have developed post-REACH as a means to instil confidence in approaches that regulators feel unable to assess. This may point to a need to increase regulatory expertise more widely in response to developing science, rather than seeking a path that thus far has stifled the application of innovation.

Continuing the theme of Brown et al.'s challenge of animal tests being predictive, as data from animal studies are extrapolated and interpreted for human relevance, this demonstrates that animal studies are also not a one-to-one replacement for humans (R6), and therefore animal tests could also be considered a form of test battery rather than having direct applicability. An illustration of this can be seen in the table of OECD test guidelines provided in the 'Landscape' report by RIVM, where rather than one single animal study TG for each toxicological endpoint there are instead multiple TGs for both in vivo and in vitro approaches.[37] However, while a change of mindset will aid the routes for acceptance of NAMs, the concerns of the biological coverage of NAMs (as discussed in the modes of action debate above) must still be addressed.

The potential of test batteries is becoming more prominent in chemical regulation, from the acceptance of the skin sensitisation DA to the development of NGRA under PARC. PrecisionTox itself is built upon a combination of assays, models, and analysis. While acceptance of these approaches is considered to take time and an expedited process to accept NAMs is sought, without meaningful engagement by regulators the current available option of WoE may become a de facto route to providing NAM data. Regulators seek to avoid case-by-case evaluations for

[37] RIVM, 'Landscape New Approach Methodologies (NAMs) for the Safety Assessment of Chemical Substances' (*National Institute for Public Health and the Environment*, June 2024) available at <https://www.rivm.nl/en/documenten/landscape-new-approach-methodologies-nams-for-safety-assessment-of-chemical-substances> accessed 22 January 2025, Table 7 at p17.

fear of time and resource burdens. However, they may find they are forced to do so, as the submission of such 'non-standard' data increases and requires consideration in REACH dossiers.[38]

3.6 CHAPTER CONCLUSION

This chapter has considered a range of risk assessment debates, discussing: their main features; the views of stakeholders; the relevance to PrecisionTox; and the policy outlook. Having observed the discussion of these identified risk assessment debates individually in scholarship and through the empirical data generated by D6.1, this critique has also served to illustrate how inextricably linked these debates are. The chosen regulatory objectives of whether to focus on hazard or exposure in risk assessment requires the consideration of data that is an imperfect substitute for human populations, whether the platform is animal testing or NAMs. Extrapolating to and recapitulation of human biology remains an area for development, but this raises policy questions as to what type and level of harms are we wish to be protected from and whether trust in claims that an exposure will not occur is sufficient to negate obtaining hazard information. Any preference for observing such hazard as a physical endpoint rather than through understanding of mechanistic information then limits efficiencies in grouping substances and screening larger numbers of substances.

Finally, with the incorporation of NAM types into test batteries allowing investigation into both hazard and exposure hypotheses, with the potential to provide rich data on a range of harmful effects and chemical mixtures, we can question whether the existing animal-based paradigm is truly protective. We can ask whether by accepting limitations of mammalian testing we can move closer to accepting that the perceived leap to NAMs as a way to strengthen regulatory risk governance may be but a few careful steps. As the need to address both the chemical backlog and substances new to the market remains, NAMs provide an opportunity to incorporate developing science. The review of the risk assessment debates suggests that a combination of factors, from acceptance of modes of action, to grouping and test batteries, can provide a protective (if not predictive) approach, regardless of hazard or exposure focus of a risk assessment, that may be scientifically, ethically, legally, and efficiently preferable to animal testing.

[38] REACH Annex VI, Step 1.

The work of PrecisionTox will assist in these aims both through the project in general as an alternative platform to animal testing, and more specifically through upcoming deliverables related to GFRB guidance and reporting templates. The solutions research of D6.3[39] and the associated study of NGO views provide further project-specific and stakeholder actions, including for example: the creation of guidance for the regulatory use of academic data; the funding of coordinated educational platforms; and developing alternative routes to validation, such as combining methods with consistent regulatory outcomes. These identified points will contribute to smoothing the debates identified in this chapter, as areas such as data availability, expertise, and acceptance are addressed.

Having considered the views offered by scholarship and stakeholder interviews to critique the current paradigm in light of potential alternatives, this report will now turn to review the judicial interpretation of legislation through an analysis of court cases. This study of case law will explore current judicial interpretation relating to the utility of chemical safety testing methods, to determine the extent to which it is considered that NAMs can be lawfully employed in chemical risk assessments processes under the status quo. In addition to the critiques of chapter 3, this case law analysis will further inform the extent and form of policy change if the uptake of NAMs is to be increased.

4. Case Law Analysis

This section provides an analysis of the EU court cases that involve different uses of chemicals governed by a wide set of EU chemical legislation. For the purposes of this analysis, we use the term 'toxicological cases' to denote a wide group of cases where the application of EU laws governing chemicals is scrutinised. The aim of this analysis is to explain the complexities associated with the judicial review of 'toxicological cases,' which is the most frequently used procedure before the EU courts with regards to chemicals. This procedure allows the Court to review the acts of different EU institutions with a remit in the field of chemicals law. Over time the list of toxicological cases has become more extensive, especially after the adoption of REACH. The full list of cases is available in Annex I.

[39] Aleksandra Čavoški, Laura Holden, Robert G Lee, 'Report on Solutions to Existing Roadblocks for Usage of NAMs in Regulation (Action Plan)' (*PrecisionTox,* 04 November 2024) <https://precisiontox.org/wp-content/uploads/2024/11/0411-D6.3-Report-on-Solutions.pdf> accessed 22 January 2025.

4.1 ARTICLE 263 TFEU AS A LEGAL AVENUE

A commonly used legal avenue in cases related to the application of EU chemicals law is the juridical review procedure, prescribed by Article 263 of the Treaty for the Functioning of the European Union (TFEU). This is the well-known action for annulment, which allows direct challenge to the legality of EU acts passed by a range of EU institutions as specified in Article 263(1). The scope of review includes legislative acts and acts of EU institutions specified by Article 263 that produce legal effects but does not extend to recommendations and opinions due to their non-binding legal nature. It also includes the "acts of bodies, offices or agencies of the Union intended to produce legal effects vis-à-vis third parties" as per Article 263(1).

The action for annulment makes a distinction between three categories of 'privileged', 'semi-privileged' and 'non-privileged' applicants. While privileged applicants do not have to prove legal interest to challenge the act, 'semi-privileged' applicants can challenge acts that fall within their operational prerogatives. This is not the case with legal and natural persons known as 'non-privileged applicants', which face very stringent conditions to meet with regards to establishing their standing.[40] Table 2 describes the types of applicants and requirements for these three types of applicants.

Table 2: Comparison of Types of Applicants Entitles to Bring an Action for Annulment

Type of the applicant	Applicant	Requirement regarding the legal interest
Privileged applicants	Member States, the European Parliament, the Council of the EU and the European Commission	They can always challenge an action without proving a legal interest
Semi-privileged applicants	The Court of Auditors, the European Central Bank and the Committee of the Regions	They can challenge the acts that fall within their remit with the aim of protecting their prerogatives

[40] See for example C-471/18P *Federal Republic of Germany v European Chemicals Agency* [2015], ECLI:EU:C:2021:48.

Non-privileged applicants	Natural or legal person	Any natural or legal person may institute the proceeding subject to the following conditions: 1) against an act addressed to that person; 2) or which is of direct and individual concern to them; and 3) against a regulatory act which is of direct concern to them and does not entail implementing measures.

The list of cases in Annex I demonstrates a steady pattern of cases initiated by legal persons who, according to EU chemical legislation, fall within one of the following categories - manufacturers, importers, and downstream users. These 'non-privileged applicants' can bring an action for annulment before the General Court. The unsuccessful applicants are allowed to submit an appeal to the Court of Justice of the EU (CJEU), which acts as the second instance court against the decision of the General Court. The decision of the CJEU is final. The question of whether a non-privileged applicant is directly and individually concerned by the decision in question applies also in toxicological cases. For example, in *European Coalition to End Animal Experiments v ECHA* case, the European Coalition to End Animal Experiments (ECEAE), which is a European animal welfare group, requested the annulment of the 'contested decision in so far as it relates to a second species pre-natal developmental toxicity study' and sought referral of the case back to the ECHA to 'consider whether there is a need to conduct a pre-natal developmental study on the registrant's substance, based on the outcome of the first study and all other relevant available data'.[41] However, the Court first had to deal with the question of admissibility by focusing on whether the ECEAE could satisfy the requirement to be regarded as an addressee of the contested decision with the prerogative to challenge the decision. After exploring the requirements prescribed by Article 263(4), the Court dismissed the action as inadmissible.

[41] Case T-673/13 *European Coalition to End Animal Experiments v European Chemicals Agency (ECHA)* [2015] ECLI:EU:T:2015:167 at para 13.

The main objective of judicial review is for the court to review the legality of the decision, which stems from Article 263(1) TFEU:

> **The Court of Justice of the European Union shall review the legality of legislative acts, of … the Commission … intended to produce legal effects vis-à-vis third parties. It shall also review the legality of acts of bodies, offices or agencies of the Union intended to produce legal effects vis-à-vis third parties.**

The Court is not concerned with the merits of the decision but focuses on its legality. Article 263(2) TFEU prescribes four main grounds of review including: lack of competence; infringement of an essential procedural requirement; infringement of the Treaties or of any rule of law relating to their application; or misuse of powers. Infringement of an essential procedural requirement and the infringement of the Treaties or of any rule of law relating to their application form the most common grounds of challenge in cases involving the application of EU chemicals law. With regards to the first category of procedural infringement, the list of cases in Annex I highlights common failures of the duty to give reasons together with breaches of the right to consultation and participation. The latter category of infringement of the Treaties or of any rule of law relating to their application most often includes the infringement of landmark legislation REACH coupled with the infringement of two key principles – the principle of proportionality and the precautionary principle.

4.2 ISSUES OF IMPORTANCE IN TOXICOLOGICAL CASES

4.2.1 Intensity of review

The intensity with which any review is conducted becomes a central issue in reviewing the legality of toxicological determinations before the court. Because the review is focussed on an expert application of toxicology, for example in classifying substances as hazardous, it is generally easier to make out procedural infringement than it is to challenge the correctness of the application of powers by an agency. A court is understandably reluctant to substitute its own view for that of the expert, but it can seek to ensure that an agency has acted throughout in accordance with its legal authority and has not acted in an *ultra vires* manner, going beyond the scope of the legal powers assigned to it.

There is some degree of deference shown by the court to the expert opinion of the agency and this governs the 'intensity' of the review. In other words, given the complexity of scientific

decision-making, some degree of discretion ought to be allowed in reaching an expert determination. This does not mean that the court will stand back from or forgo review. This is because the court is charged in Article 263(1) TFEU with testing the legality of the decision made, including difficult questions of whether the powers of the decision-maker were appropriately employed. In a case challenging the withdrawal of authorisations for plant protection products by the European Commission, the judicial task was described as follows:

> **If the Commission is to be able to pursue effectively the objective assigned to it, account being taken of the complex technical assessments which it must undertake, it must be recognised as enjoying a broad discretion. However, the exercise of that discretion is not excluded from review by the Court. The Court has consistently held that in the context of such a review the Community judicature must verify whether the relevant procedural rules have been complied with, whether the facts admitted by the Commission have been accurately stated and whether there has been a manifest error of appraisal or a misuse of powers**.[42]

4.2.2 Standard of review

As Craig and De Burca point out, judicial review 'entails challenge to law, fact, and discretion.'[43] However, the standard of judicial review is likely to vary across these parameters and the Court is tasked with examining 'whether the exercise of the discretion was vitiated by a manifest error, misuse of power, or clear excess in the bounds of discretion'.[44] The manifest errors of assessment of properties of chemicals in question is a common theme across all of the toxicological cases listed in Annex I. This includes the review of assessment of highly complex scientific and technical facts, which subsequently inform the decision that EU institutions (most commonly ECHA) need to take. With regards to chemicals deployed on the EU market, the European Commission is the ultimate risk manager of a complex process of chemical safety assessment that involves ECHA and member states. It is worth pointing out that much of the work regarding all phases of evaluation, authorisation, and restriction of chemicals involves ECHA, though member states have an important role to play, in particular in substance evaluation.

---

[42] Case C-326/05 P *Industrias Químicas del Vallés SA v Commission of the European Communities* [2007] ECLI:EU:C: 2007:443 at paras 75 and 76.

[43] Paul Craig and Gráinne de Búrca, *EU law: text, cases, and materials*, Oxford University Press, 2020 at p. 631

[44] Ibid.

As might be expected, the extent of discretion is even greater in areas 'of evolving and complex technology' where EU institutions are faced with the assessment of 'highly complex scientific and technical facts, in order to determine the nature and scope of the measures which they adopt.'[45] The main objective of review is to ensure that the EU institutions that have passed the measure 'actually exercised their discretion, which presupposes that they took into consideration all the relevant factors and circumstances of the situation the act was intended to regulate'.[46] Due to this wide discretion in assessing complex and technical cases, the role of the EU courts has to remain limited whereby the EU adjudicator 'cannot substitute its assessment of scientific and technical facts for that of the authorities of the European Union on which alone the FEU Treaty has placed that task.'[47] Thus, the discussion in cases often revolves around the 'reliable science' in the form of scientific studies presented by both parties in order to advance or negate the hazardous nature of the chemicals under scrutiny and to assert the weight of evidence to be applied.

There are several good illustrations. In *Global Silicones Council and Others v European Commission* case, the first plea contended manifest error in the assessment of properties of three chemical substances: octamethylcyclotetrasiloxane ('D4'), decamethylcyclopentasiloxane ('D5') and dodecamethylcyclohexsiloxane ('D6') individually or in a mixture. The applicants made a reference to a study by Krogseth et al. from 2017, which suggested that the substances in question do not display 'trophic magnification'.[48] Although the Court found that "the data resulting from the field studies indicating that there was no trophic magnification in certain food webs were taken into consideration", the Court referred to the findings of the ECHA Member State Committee ('the MSC').[49] The MSC's finding shows that the picture is more complex as even if a substance does not bioaccumulate by means of biomagnification, this may occur through means of bioconcentration.[50] As the MSC concluded,

[45] Cases T-134/13 *Polynt SpA and Sitre v European Chemicals Agency* (ECHA) [2015] ECLI:EU:T:2015:254 at para 52 and T-226/18, *Global Silicones Council and Others v European Commission* [2021] ECLI:EU:T:2021:403 at para 74 (There is currently an appeal before the Court of Justice of the EU - Case C-558/21 P).

[46] See cases C-343/09 *Afton Chemical* [2010] EU:C:2010:419, paras 33 and 34, T-689/13 *Bilbaína de Alquitranes and Others* v *European Chemicals Agency* [2015] ECLI:EU:T:2015:767 para 77 and T-134/13 *Polynt SpA and Sitre v European Chemicals Agency* [2015] ECLI:EU:T:2015:254 at para 53.

[47] Case T-134/13 *Polynt SpA and Sitre v European Chemicals Agency* (ECHA) [2015] ECLI:EU:T:2015:254 at para 52 and T-226/18, *Global Silicones Council and Others v European Commission* [2021] ECLI:EU:T:2021:403 at para 74.

[48] Case T-519/18 *Global Silicones Council and Others v European Chemicals Agency* [2021] ECLI:EU:T:2021:404 at para 66.

[49] Ibid at para 74.

[50] Ibid.

'the absence of biomagnification of a substance in a food chain does not prove that that substance does not biomagnify in other food chains.'[51]

Equally, in *PlasticsEurope v ECHA* case, the appellant argued that ECHA based its analysis on the Lee study while disregarding various other high-quality studies, 'such as Goodman (2009, 2006), Gray (2004), Center for the Evaluation of Risks to Human Reproduction (2008), EFSA (2015) and SCOEL (2014)'.[52] In the *CWS Powder Coatings GmbH v European Commission* case, the parties argued about the reliability of the *Heinrich* study.[53] The applicants challenged the reliability and acceptability of data obtained from the *Heinrich* study,[54] which in their view formed the basis of the decision of ECHA's Committee for Risk Assessment ('the RAC') on the classification and labelling of titanium dioxide (as carcinogenic to humans by inhalation in powder form). Moreover, the applicant stressed that the authority in France (the Agence nationale de sécurité sanitaire de l'alimentation, de l'environnement et du travail (ANSES)), who initiated the new classification and labelling found this study to be unreliable as it was conducted solely on female rats and it was based on the administration of a single excessive testing dose.[55] On the other hand, the European Commission argued that it deployed other studies, such as the *Lee* study,[56] which informed their findings.[57]

In such cases, there is considerable discussion as to whether all existing, reliable and measurable data are taken into account so as to demonstrate scientific excellence in chemical risk assessment.[58] This needs to be put into a wider context whereby decision-makers are often faced with cases where scientific knowledge may not be fully conclusive. As a result, decision-makers have to decide on partial or incomplete studies, which leads to instances in which perfect evidence is not always available. This then puts an onus on the Court, which has a task 'to establish, in the light of the factors relied on by the applicant, whether the evidence relied on is factually accurate, reliable and consistent, whether that evidence contains all the information

[51] Ibid.
[52] C-876/19P *PlasticsEurope v European Chemicals Agency* [2021]ECLI:EU:C:2021:1047 at para 47.
[53] See Cases T-279/20 and T-288/20 *CWS Powder Coatings GmbH and Others v European Commission* [2022] ECLI:EU:T:2022:725.
[54] Heinrich U, et al. Chronic inhalation exposure of Wistar rats and two different strains of mice to diesel exhaust, carbon black and titanium dioxide 1995;7 *Inhalation Toxicology* 533–556.
[55] Cases T-279/20 and T-288/20 *CWS Powder Coatings GmbH and Others v European* at para 50.
[56] K.P. Lee, et al. Pulmonary response to impaired lung clearance in rats following excessive TiO2 dust deposition, 1986 41(1) *Environmental Research*, 144-167.
[57] See more in Čavoški, A., Holden, L. & Lee, R., Reviewing science-based decisions: CWS Powder Coatings GmbH v European Commission, Environmental Liability - Law, Policy and Practice. 28, 1, p. 16–22.
[58] E.g. Case T-636/19 *Chemours Netherlands v European Chemicals Agency* [2022] ECLI:EU:T:2022:86.

which must be taken into account in order to assess a complex situation, and whether it is capable of substantiating the conclusions drawn from it'.[59] Thus, it is not surprising that contestation concerning what constitutes reliable evidence remains the cornerstone of the majority of the cases of substantive (rather than procedural) legal challenge.

The weight of evidence also plays an important role in assessing the regulatory relevance and reliability of scientific studies. This is even more the case with regards to alternatives to animal testing, where additional insights into biological mechanisms could significantly add to the weight of evidence.[60] The case law related to chemicals provides some clarifications regarding several terms surrounding the concept of weight of evidence. In *PlasticsEurope v European Chemicals Agency* case, the General Court suggested that 'expressions "scientific evidence" and "scientific knowledge" are synonyms.[61] The status of scientific knowledge or scientific methods or rules can only be recognised where those elements are based on scientific evidence'.[62] Because the assessment is seeking to prove 'hazard', it was emphasised that scientific findings should be based on the 'possible' undesirable effects of a substance, not necessarily its 'probable' effects'.[63] According to the Court, in exercising the weight of evidence, ECHA is bound by the principle of scientific excellence in analysing the intrinsic properties of a substance.[64] *Sasol Germany and Others v Commission* case further expanded on weight of evidence approaches by ruling that this concept infers compliance with the best current scientific standards.[65] However, the Court specified that case law does not suppose that every assessment should follow a 'specific and uniform methodological approach', though there is no doubt that weight of evidence has to be applied in all assessments.[66]

Another important consideration regarding weight of evidence can be identified from the current case law governing chemicals in the EU. This is the question of consequence placed on data and what the weight of evidence determination involves.[67] In particular, both quantitative and

[59] See Cases C-525/04 P, *Spain* v *Lenzing* [2007] EU:C:2007:698 at para 57; C-405/07 P, *Netherlands* v *Commission* [2008] EU:C:2008:613 para 55; T-257/07 *France* v *Commission* [2007] EU:T:2011:444 para 87 and T-636/19 *Chemours Netherlands v European Chemicals Agency* [2022] EU:T:2022:86 para 48.
[60] See Case T-636/17 *PlasticsEurope v European Chemicals Agency* [2019] ECLI:EU:T:2019:639 at paras 93-94. This decision is now being appealed before the CJEU.
[61] Ibid at para 93.
[62] Ibid.
[63] Ibid at para 98.
[64] Ibid at para 94.
[65] Case T-661/19 *Sasol Germany and Others v Commission* [2021] ECLI:EU:T:2021:779 at para 35.
[66] Ibid at para 35.
[67] See Case T-226/18, *Global Silicones Council and Others v European Commission* [2021] ECLI:EU:T:2021:403.

qualitative methods may be available, raising the question of what weight is given to these different approaches. In *Global Silicones Council and Others v European Commission* case, the applicant questioned the credibility of a weight of evidence determination due to the fact that no quantitative weight was apportioned to each piece or body of evidence.[68] The applicant also argued that it was necessary to deploy a quantitative approach for reasons of transparency and to prevent arbitrary decisions.[69] The Court analysed the linguistic interpretation of the term 'appropriate weight' as stipulated by Annex XIII of REACH, as well as the text of the ECHA's 'Guidance on Information Requirements and Chemical Safety Assessment'. To give appropriate weight to the quality and consistency of the data, the Court concluded that, according to rules set out in Annex XIII, the competent authorities can apply both quantitative and qualitative approaches. The choice between the quantitative or qualitative weight of evidence determinations may vary depending on the circumstances of each case, which itself may rest upon the nature of the available information.[70]

4.3 ANIMAL WELFARE

Animal welfare concerns have been raised in numerous toxicological cases over the years. These entail questions about the efficacy of animal testing, often coupled with wider questions about the suitability of mammalian species to provide a holistic picture of toxicity and their human relevance. Equally, the need to conduct animal testing when there is available alternative evidence features prominently in the caselaw. This is of particular significance as REACH commits to deploying animal testing only as a last resort.[71]

Despite this commitment, the deployment of NAMs still remains limited and the adaptations prescribed by Annex XI of REACH have gained relatively little traction. Toxicological cases shed some light on the reasoning deployed both by the regulator and the EU courts. It is very often the case that as a result of power vested to ECHA, Member States, and the European Commission under REACH, in particular Chapters 2 (Dossier Evaluation) and 3 (Substance Evaluation), the applicants (mostly industry) were asked to provide additional information by conducting some form of further testing. This usually requires additional testing on animals, which in most cases is defended by claims of the need to gather comprehensive data to assess the safety of

[68] Ibid at para 147.
[69] Ibid.
[70] Ibid at para 149.
[71] Article 25(1) REACH.

chemicals in question.[72] The arguments of applicants are in most cases multifaceted. It is argued that additional testing is unnecessary as existing studies are sufficient, and that the case for further testing lacks solid scientific foundation. These arguments are supported by appeals to ethical considerations in scientific research and the need to better use alternative testing results submitted by applicants relying on adaptations. For example, in *Polynt SpA's v European Chemicals Agency* case the applicant argued without success that conducting the EOGRTS would result in the death of approximately 600 animals, violating both the company's policies and broader animal welfare principles enshrined in Article 13 TFEU and Article 25(1) REACH, which states that animal testing should be a last resort.[73]

Finally, the question of human relevance of species used for animal testing is far from academic. This is an issue that was raised before the courts. In *Deza, a.s. v European Commission* case, the limitations of use of rats and mice for certain types of testing was emphasised.[74] For example, the OECD guidance document makes a reference to the limitations of mammalian studies in certain circumstances depending on the strain of these species and, for example, their unsuitability for studying the carcinogenicity of substances due to their high predisposition to the growth of tumours.[75] However, the outcome in most cases has not necessarily been conducive to less testing because the courts, in agreement with the regulator's view, see further testing as a means of ensuring safety for humans and the environment. The same conclusion can be drawn from the analysis of the Board of Appeal hearings, the remit of this is to review appeals 'for any natural or legal person affected by decisions taken by the Agency'.[76]

Animal welfare considerations are raised in relation to other relevant EU legislation governing the use of chemicals. This issue is closely associated with the Regulation on the Classification, Labelling and Packaging of Substances and Mixtures (CLP),[77] as the other key legal instrument in this space. Despite stipulating certain non-animal methods in Annex 1, CLP is well known for its preference for animal test data due to its endpoint-specific sections in Part 3 of Annex 1. It is not surprising then that animal testing often forms the basis of the decision on classification and

[72] See for example C-471/18/P *Germany v Esso Raffinage* [2021] ECLI:EU:C:2021:48; T-125/17 *BASF Grenzach GmbH v European Chemicals Agency* [2018] ECLI:EU:T:2019:638; T-655/20 *Symrise v European Chemicals Agency* [2021] ECLI:EU:T:2023:736; T-207/21 *Polynt SpA v European Chemicals Agency* [2023] ECLI:EU:T:2023:361; T-868/19 *Nouryon Industrial Chemicals and Others v* Commission [2020] ECLI:EU:T:2023:168.

[73] Case T-207/21R *Polynt SpA v European Chemicals Agency* [2023] ECLI:EU:T:2023:361 at paras 106-107.

[74] Case C-813/18 P *Deza, a.s. v European Commission* [2020]ECLI:EU:C:2020:832 at paras 73-75.

[75] Ibid at paras 75 and 76.

[76] Recital 106 REACH. For ECHA Board of Appeals cases, see https://echa.europa.eu/regulations/appeals.

[77] Regulation (EC) No 1272/2008 of the European Parliament and of the Council of 16 December 2008 on Classification, Labelling and Packaging of substances and mixtures (CLP) [2008] *OJ L 353/1*.

those decisions are challenged for their reliance upon unreliable and out-dated animal testing that does not necessarily prove a causal link between harm and the use of a substance.[78]

The analysis of the toxicological cases also reveals the tension between the Cosmetics Regulation and REACH, since regulation under the latter may demand animal testing of ingredients in cosmetics whereas the former seeks to prohibit mammalian testing of cosmetic products. This leads to some confusion for the general public, which is under the misapprehension that all substances used in cometic products are free from animal testing.[79] Several of these issues have been raised in caselaw in relation to cosmetics and cosmetic substances. The potential breach of the Cosmetics Regulation that would result for the animal testing requirement have been raised by the parties, coupled with the need to better use existing data.[80] Industry applicants also raised concerns about the impact that animal testing may have on their reputation and market position.[81] In *Symrise v European Chemicals Agency* case, the Court emphasised that while the Cosmetics Regulation may ban animal testing for consumer safety, REACH can still require animal testing to address other safety issues attaching to ingredients in cosmetics, such as environmental protection or occupational safety.[82]

This is not to say that the Court is entirely ignoring the need to consider animal welfare at the same time in the light of the 3Rs commitment to replace, reduce, or refine animal testing. In the *Federal Republic of Germany v Esso Raffinage* case, the CJEU has discussed the use of animal testing in the light of the 3Rs commitment. The Court decided that:

> **A registrant has, generally and therefore especially where ECHA issues it with a decision asking it to complete its registration dossier with a study involving animal testing, not simply the possibility but the obligation to generate information obtained by means other than animal testing 'whenever possible' and to undertake such testing 'only as a last resort**.[83]

[78] See for example Caase T-400/17 *Deza, a.s. v European Commission* where Deza challenged the classified anthraquinone as a substance with presumed carcinogenic potential for humans.
[79] See for example Cases T-226/18 *Global Silicones Council and Others v European Commission* [2021] ECLI:EU:T:2021:403; T-176/19 *3V Sigma SpA v European Chemicals Agency* [2020]ECLI:EU:T:2020:621; T-125/17 *BASF Grenzach GmbH v European Chemicals Agency* [2019]ECLI:EU:T:2019:638.
[80] See Case T-125/17 *BASF Grenzach GmbH v European Chemicals Agency* [2017] ECLI:EU:T:2017:496 and Case T-655/20 *Symrise v European Chemicals Agency* [2023] ECLI:EU:T:2023:736.
[81] Case T-655/20 *Symrise v European Chemicals Agency* [2023] ECLI:EU:T:2023:736.
[82] Ibid at paras 89-114.
[83] Case C-471/18 P *Federal Republic of Germany v Esso Raffinage* [2021] ECLI:EU:C:2021:48 at para 132.

In *Polynt SpA v ECHA* case, the Court recognised the importance of minimising the animal testing in the light of Article 25 REACH but found that ECHA reached an appropriate balance between animal welfare considerations and ensuring chemical safety.[84] Yet endorsements by the Court of requirements for additional animal testing inevitably call into question whether the 'last resort' principle is being applied.

4.4 ROLE OF LEGAL PRINCIPLES IN PURSUING 3RS COMMITMENT

Behind this issue of the oversight of the last resort requirement is the role of legal principles in assessing the need for additional animal testing. In several cases, the Court applied wider legal principles as a tool to resolve challenges brought by the applicants concerning the need to undertake animal testing, where they asserted that regulatory demands for animal testing went against this commitment of deploying such testing only as a last resort. The case for animal testing was assessed through the application of key principles of EU law, in particular the precautionary principle, the principle of high level of protection of human health and the environment, the principle of proportionality, and the principle of legitimate expectation. The precautionary principle plays a prominent role in toxicological cases as it is used to demonstrate the value of additional testing, as explained above, to ensure comprehensive risk assessment. As espoused by the Principle 15 of the Rio Declaration, the principle states that: "Where there are threats of serious or irreversible damage, lack of full scientific certainty shall not be used as a reason for postponing cost-effective measures to prevent environmental degradation".[85] This might be read as suggesting that in the absence of certainty, further testing, including animal testing ought not to be postponed. Although the European Commission states its commitment to the 3Rs principle, the precautionary principle may allow a discretion to the regulator in balancing the need for further testing as opposed to ensuring animal welfare.

In one of the recent cases, *Polynt SpA v European Chemicals Agency* case,[86] the Court reflected on two key EU principles – the principle of proportionality and the principle of ensuring a high level of protection of human health and the environment.[87] The former principle requires careful

[84] Case T-207/21 *Polynt SpA v European Chemicals Agency* [2023] ECLI:EU:T:2023:361.
[85] 31 ILM 874 [1992] Available at: <https://www.un.org/en/development/desa/population/migration/generalassembly/docs/globalcompact/A_CONF.151_26_Vol.I_Declaration.pdf> last accessed 20 January 2025.
[86] Case T-207/21 *Polynt SpA v European Chemicals Agency* [2023] ECLI:EU:T:2023:361.
[87] Article 5(4) TEU.

consideration of measures demanded in order to avoid imposing onerous and disproportionate burdens when weighed against the aim to be achieved.[88] This in the case of animal testing would entail requiring less testing by placing greater reliance on existing data on the chemicals being tested, for example by grouping and read across. However, this needs to be balanced by ensuring, at the same time, a high level of protection of human health and the environment, through a thorough assessment of the risks involved. As stated in C-558/07, in order to achieve that objective, "as recital 19 of the REACH Regulation states, the registration obligation imposed on manufacturers and importers, which includes the obligation to generate data on the substances which they manufacture or import, to use those data to assess the risks related to those substances and to develop and recommend appropriate risk-management measures".[89] As a result, the Court confirmed that in certain situations the information requirements set out in Annexes VII to X to the REACH Regulation will inevitably entail animal testing as "only testing on vertebrate animals will provide sufficient scientific information to enable measures to be taken to protect human health and the environment".[90] This statement may be seen as somewhat contradictory to the overall objective of deploying animal testing as a last resort without providing more clarity as to when animal testing would be the only option that provides comprehensive assurance.

4.5 CHAPTER CONCLUSION

This brief analysis of the court cases addressing the application of the EU chemicals law reveal significant complexities for all actors involved in the decision-making process. It may be argued that the key challenge comes from the fact that toxicological cases involve complex science where decision-makers are faced with often incomplete or conflicting evidence, which informs their decision-making. This, coupled with the extensive discretion involved in scientific assessment, means it is not therefore surprising that manifest error of assessment is questioned in each case. The other challenge comes from stringent legal requirements stipulated in law governing chemicals, such as REACH, which are instituted to ensure high levels of protection of human health and environment. This opens up a question of extent to which interpretation of these requirements by the regulator may inhibit the use of animal testing and thus lead to challenges in court. Finally, these cases expose high levels of complexity of the assessment required, the scientific competence of the decision makers, and the wider political

[88] Case C-15/10, *Etimine SA v Secretary of State for Work and Pensions* [2011] EU:C:2011:504 at para 124.
[89] Case C-558/07 *S.P.C.M. and Others* [2009] EU:C:2009:430 at paras 45 and 46.
[90] Case T-207/21 *Polynt SpA v European Chemicals Agency* [2023] ECLI:EU:T:2023:361 at para 108.

accountability of the decision-makers involved, which further demonstrates the difficulties surrounding judicial review of scientific methodology behind regulatory determinations.

Having considered different conceptual approaches to risk assessment in chapter three and the current application of law by the courts in this chapter, the final chapter turns to a doctrinal study of legislation, to determine whether there is the capacity for existing legal text to be reinterpreted to allow the application of NAMs.

5. Capacity for the Use of NAMs in Legal Frameworks

With contributions from Louis Dawson

This report has considered the current state-of-play of animal uses in science and analysed key debates relating to the challenge of the existing paradigm from alternative approaches. Having reviewed relevant legal cases relating to submissions of NAM data and the judicial interpretation of legislation, this final chapter provides an analysis of key legislative texts in the chemical domain to uncover the extent that NAMs are recognised in EU regulatory structures. The judicial interpretation of legislation as described in the preceding chapter provides a 'law in action' interpretation, as construed by societal norms and particularly in acknowledgment of the authority and expertise of regulatory agencies. This chapter turns to focus on a 'letter of the law' study to interpret what may be possible as discerned from the direct wording of the legal text.

This analysis will provide an assessment as to the possible legal conformity of NAMs with EU law on chemical risk assessment, which is based on a mapping exercise of chemical risk assessment in EU legislation across the following sectors: industrial chemicals; classification, labelling and packaging; cosmetics; plant protection products; biocidal products; food contact materials; and pharmaceuticals.[91] All of these have traditionally incorporated a level of toxicological testing to determine hazard or risk and thus suitability for market, with pharmaceuticals included to address the cross-agency 'One Health' policy.[92] An overview of key attributes of each legislative instrument is provided at Annex II, which indicates similarities and differences in their remit, risk governance steps, actor roles, and endpoint prioritisation. This table highlights key cross-mentions of other pieces of legislation (particularly of note is the common thread of CLP) and the

[91] Legislation (n2).

[92] ECDC, 'Cross-agency One Health Tast Force Framework for Action 2024-2026' (*European Union,* 2024) available at:<https://www.ecdc.europa.eu/sites/default/files/documents/cross-agency-one-health.pdf> accessed 22 January 2025.

endpoints of particular concern, which can inform future NAM development. The legislation tends to focus on substances, mixtures and articles, though in some cases, such as cosmetics and biocidal products, there is a focus on the final formulation/product. Except for the CLP, there is an expectation for a full risk assessment to be undertaken – including hazard characterisation and exposure assessment – and this is regardless of whether pre-market authorisation is needed or not. A varied approach to regulating chemicals can be observed: for industrial chemicals and CLP, the applicant/registrant is responsible for carrying out risk assessment as part of their application, with regulatory agencies and the European Commission acting as scrutineers, providing an oversight function and focusing on union harmonisation. Much of the remaining legislation, however, while requiring applicants to meet data requirements, also confers an approval role upon regulators prior to marketing.

A table of comparing toxicological and ecotoxicological endpoints by legislation is included at Annex III. A large variety of endpoints to be identified can be observed, the most common being serious eye damage and irritation, skin corrosion and irritation, skin sensitisation, and acute toxicity. The remaining endpoints are addressed to varying degrees across the selected legislation, with some decided on a case-by-case basis depending on the normal and reasonably foreseeable use of the product. This is particularly the case for cosmetics, which has purposely been left blank, since endpoints are determined by 'relevance', and are not fixed *per se*. Similarly, although much fuller, BPR refers to the 'core data set' only, and thus, as is the case with the whole table, represents the bare minimum requirements. Other than for food contact materials (FCM), all endpoints refer directly to the associated legislation. FCM makes reference to associated guidance, which can be considered as being given legal force by the framework regulation; plastics and active materials,[93] each of which have specific measures, are used as exemplars. In many cases, differences can also be observed in how endpoints are separated or grouped, for example mutagenicity, germ cell mutagenicity, and genotoxicity. The effects of terminological differences on resulting interpretation also become apparent through the tabular formatting, in that the CLP (for which there is a limited testing requirement), is alone in stating endpoints of 'specific target organ toxicity – single exposure' and for 'aspiration hazard.'

The overarching Directive relating to animal use is Directive 2010/63 on the protection of animals used for scientific purposes.[94] Here, the definition of an animal is provided as being live, non-

[93] Active and intelligent materials enhance foods or give information about food condition. For more see EFSA, 'Active and Intelligent Materials' (*European Union,* 19 September 2024) available at <https://www.efsa.europa.eu/en/topics/active-and-intelligent-materials> accessed 24 January 2025.
[94] Directive 2010/63 on the Protection of Animals Used for Scientific Purposes [2010] OJ L 276/33 ('2010/61').

human vertebrate animals (including independently feeding larval forms and foetal forms of mammals as from the last third of their normal development), and live cephalopods.[95] The expectation on Member States from the Directive is that 'wherever possible, a scientifically satisfactory method or testing strategy, not entailing the use of live animals, shall be used instead of a procedure.'[96] This clearly allows for the NAMs to be used in place of animal studies. The rest of this chapter reviews the text of each selected regulation individually, to determine the extent the Directive can be implemented.

5.1 INDUSTRIAL CHEMICALS: REACH

Within REACH, Article 13(1) states:

> **Information on intrinsic properties of substances may be generated by means other than tests, provided that the conditions set out in Annex XI are met. In particular for human toxicity, information shall be generated whenever possible by means other than vertebrate animal tests, through the use of alternative methods…**

Therefore, animal tests on intrinsic properties do not have to be conducted, but Annex XI must then be followed. The term 'intrinsic properties' is not defined within the legislation, but is generally taken to mean the lists of endpoints in Annexes VII-X. The 'shall' imperative mandates that alternative methods rather than vertebrate animals are to be used whenever possible. This Article goes on to mention methods such as 'In vitro' by way of example, so that other methods are possible.[97] Amending the Test Method Regulation on adopted methods is,[98] however, a long process, and it is apparent that the OECD Test Guideline (TG) route is used as a proxy pending update of the TMR, with TGs representing an international test method 'recognised by the Commission or [ECHA] as being appropriate.'[99]

Annex VI provides the steps on gathering information to fulfil requirements. The registrant should not disregard information on hazardous endpoints just because they are not within the applicable tonnage band, and it seems that alternative methods are acceptable. There is no mention of the need for validation, although the text states that information is to 'assist' in what appears to be the start of a weight of evidence (WoE) approach. That exposure is also to be

[95] 2010/63 Article 1(3).
[96] 2010/63 Article 4(1).
[97] REACH Article 13(1).
[98] REACH Article 13(2).
[99] REACH Article 13(3).

considered introduces the possibility for exposure-based waiving. Under step 3, rather than validation the requirement is that existing information has 'relevance' and is of 'sufficient quality.' This may be necessary, however, because not all existing animal tests have been validated. When carrying out tests / methods in step 4, the order of approaches to be employed is set as 'other data sources' first and then 'new tests on vertebrates.'

Annex XI is key for adaptations from animal tests. Adaptations seem open to ECHA's interpretation and acceptance at dossier evaluation, as it is ECHA who 'may assess' these adaptations.[100] While in vitro methods only waive animal studies where the former are validated,[101] there remains other options for the use of existing data through grouping and read-across and developed QSAR models with 'scientific validity'.[102] Information under a WoE approach,[103] however, seems the most likely path for utilising new data from NAMs and there is neither a stated need for validation nor particular TMR requirements nor need for international acceptance of method. Indeed, reference to WoE here includes the statement that 'there may also be sufficient Weight of Evidence from the use of newly developed test methods, not yet included in the [Test Methods Regulation], leading to a reasoned justification that they provide the information that would enable a conclusion on the information requirement.' WoE requires that more than one alternative method would need to be used, which might consist of the combination of methods demonstrating different key events of an adverse outcome pathway.

Yet difficulty arises by requiring that the justification of a WoE approach should address the information that would have otherwise been obtained from the animal test. There is no clear or definitive guidance on exactly what evidence is expected to be provided by NAMs. We propose that, if the NAMs are an improvement on the animal test's relevance, this can form part of the justification. This is because of the wording that WoE 'must have regard to' rather than requiring it to be the same information that would have been obtained from the otherwise stated animal study. Dossier evaluations and Board of Appeal cases indicate it is information such as observations of organ weights and offspring behaviour that is being sought.[104] However, this is

[100] REACH Annex XI.
[101] REACH Annex XI s1.4. It is worth noting that validation to 'internationally agreed validation principles' infers OECD, but in theory could be in adherence to another process, such as ISO.
[102] REACH Annex XI s1.5 and 1.3, respectively.
[103] REACH Annex XI s1.2.
[104] Dossier evaluation of substance EC number 271-234-0, Para 52: "Relevant information that can be used to support weight of evidence adaptation for the information requirement of Annex IX, Section 8.6.2 includes similar information that is produced by the OECD TG 408 [90-day oral toxicity study]… requires the study to investigate the following key elements: A) in-life observations, B) blood chemistry, C) organ and tissue toxicity." Available at <https://echa.europa.eu/documents/10162/1ac7f4f3-f05e-c086-e612-19fcc4dbd14e> accessed 22 January 2025. The Board of Appeal's 'Digest of Decisions' para 11.3.9 expects WoE to meet "the information

nearly impossible to provide as NAMs often identify indicators of earlier onset effects or of bioactivity. This could be overcome by a change in interpretation by regulators and the BoA, to the effect that the information to be obtained relates to harmful effects, for example as part of an AOP.[105]

5.2 Classification, Labelling, and Packaging (CLP)

Article 9 of CLP states that when evaluating hazard information for substances and mixtures, where criteria cannot be applied to available information, an evaluation is to be carried out ('shall') by applying 'a weight of evidence determination using expert judgment.' Annex I of CLP goes on to give examples including 'suitable in vitro tests' and studies about 'site of action and mechanism or mode of action.'[106]

While the preamble of CLP states a preference that information is generated in accordance with methods such as those under REACH or international principles or procedures for validation,[107] this is not explicit within the body of the legislation and WoE is a method available in REACH, just as it is in CLP. The preamble goes on to state:

> **This Regulation should take the utmost account of promoting alternative methods for the assessment of hazards of substances and mixtures and of the obligation to generate information on intrinsic properties by means other than tests on animals... Future criteria should not become a barrier to this aim and the corresponding obligations under that Regulation, and should under no circumstances lead to the use of animal tests where alternative tests are adequate for the purposes of classification and labelling.**[108]

requirements for the respective endpoint, e.g. the key parameters need to be covered...[for an extended one-generation reproductive toxicity study (EOGRTS)] a registrant must demonstrate that the available information adequately identifies and characterises the pre-natal developmental toxicity of the substance at issue"; under para 11.2.5, an EOGRTS is expected to show "at least one of the following: (1) adverse effects on reproductive organs, (2) adverse effects on reproductive tissues, or (3) other concerns in relation to reproductive toxicity." Available at <https://echa.europa.eu/documents/10162/2314761/digest_of_decisions_of_boa_en.pdf/cad5c04e-1888-9ac3-5718-eb6f17a395a8?t=1642149879775> accessed 22 January 2025. 11.2.5

[105] This section has provided a summary doctrinal analysis. Full legislative mapping of REACH is available at Annex IV of this report.

[106] CLP Annex I, s1.1.1.3.

[107] CLP Preamble note 21.

[108] CLP Preamble note 27.

It can be taken, therefore, that new animal studies are not in the spirit of the aims of the CLP, and while validation may be a preferred route, the opportunity to utilise WoE subsists.

In light of the UN's Globally Harmonised System (GHS) upon which CLP is based,[109] in the detail of endpoint classification the GHS makes strong reference to animal tests, which CLP transposes. However, parameter 2 of GHS is that it is not based on uniform test methods and as such it is test-method neutral and not tied to OECD TGs.[110] While GHS appears to require validated tests,[111] mention of WoE only refers to 'valid in vitro' as an example,[112] inferring that some level of acceptance criteria is expected. However, the GHS carries no legal force, and the CLP has diverged from the GHS, most recently in the inclusion of new endpoints such as endocrine disruption.[113]

5.3 PLANT PROTECTION PRODUCTS (PPP)

The testing for PPP is to be under GLP, using validated alternatives that are OECD test methods and guidance documents. A list of data requirements with the test methods and guidance for PPP is provided via an EC communication.[114] The avoidance of using animals is acknowledged to be from an ethical point of view, however, reference such as for using the oral route of administration implies animal tests. Acute toxicity also demands details on behavioural changes, clinical signs, and gross pathology, all of which are animal-based parameters. For acute toxicity, reference to alternative methods directs to justifications under CLP.[115]

In addition to CLP justifications, a weight of evidence using tiered testing is also suggested for skin and eye irritation, starting with in vitro and progressing to in vivo, plus a second species. Animal tests are otherwise specified for skin sensitisation (OECD guinea pig); short term toxicity (rat and dog); long term toxicity (rat and mouse), and reproductive toxicity (rat and

[109] UN, 'Globally Harmonised System of Classification and Labelling of Chemicals (GHS)' (*United Nations Economic Commission for Europe,* 2023) available at <https://unece.org/sites/default/files/2023-07/GHS%20Rev10e.pdf> accessed 22 January 2025.
[110] Ibid, 2(i). Although it is worth noting that classifications do then refer to specific tests, for example 3.1.2.3: 'the rat or rabbit are preferred for evaluation of acute dermal toxicity', indicating a point of contradiction.
[111] Ibid, 1.3.2.4.2.
[112] Ibid, 1.3.2.4.9.
[113] For more, see: UN, 'Proposal to Reconsider the Inclusion of Endocrine Disruptor in the GHS Hazard Classification' (*United Nations Economic Commission for Europe,* 20 November 2023) available at <https://unece.org/sites/default/files/2023-11/UN-SCEGHS-45-INF15e.pdf> accessed 22 January 2025.
[114] Commission Communication in the Framework of the Implementation of Commission Regulation (EU) No 284/2013 Setting Out the Data Requirements for Plant Protection Products, in Accordance with Regulation (EC) No 1107/2009 Concerning the Placing of Plant Protection Products on the Market (2013/C 95/02) 03 April 2013.
[115] PPP Annex, Part A, Section 7.1.1 - 7.1.6.

rabbit). Feeding studies and studies to understand the metabolism, distribution, and expression in laying hens, lactating goats and cows, and fish are also required, and these are to aid dietary risk assessment relating to transfer from animal feed.

The Implementing Regulation on principles for PPP authorisation does state that while models are to be reliably validated, if they have not been they should 'be supported with details indicating how the model calculates estimates provided, and explanations of all the inputs to the model and details of how they have been derived.' [116] While this appears beneficial for NAMs, this regulation also states that the impact on health is to include the acceptable operator exposure level (AOEL) in milligrams of the chemical per kilogram body weight of the operator based on 'tests in the most sensitive relevant animal species' or appropriate, available data from humans.

The PPP regulation is therefore very focused on animal testing in the reality of its requirements and guiding principles.

5.4 BIOCIDAL PRODUCTS REGULATION (BPR)

Article 6 of BPR provides that an application for approval of an active substance shall include a dossier, satisfying the requirements set out in Annex II, which affirms the 'importance of reducing testing on vertebrates' and provides that 'new tests' involving vertebrates are only to 'be conducted as the last available option to comply with the data requirements [...] when all the other data sources have been exhausted'. To facilitate this, Annex II requires the applicant to initiate a pre-submission consultation, with a particular emphasis on vertebrate testing. BPR permits applicants to 'not provide' dossier data where either (a) the data is not necessary owing to exposure associated with the proposed uses; (b) it is not scientifically necessary to supply the data; or (c) it is not technically possible to generate the data.[117]

In setting the data requirements, the table at Annex II of the BPR provides 'specific indications' for the adaption of data elements to reduce vertebrate testing. Given that the provision aims to reduce (rather than, say, replace) such testing, the table merely references the adaption of 'some', not all, data elements; for example, 'acute toxicity', 'by oral', 'by inhalation' and 'by

[116] Commission Regulation (EU) No 546/2011 Implementing Regulation (EC) No 1107/2009 as Regards Uniform Principles for Evaluation and Authorisation of Plant Protect Productions [2011] OJ L 155/127, Annex 2.6(e) and Annex Part A 1.4.1.1(a)(i).
[117] BPR Article 6(2).

dermal' route data requirements (depending on the route of exposure to humans), contain no, or very limited, indications of adaptions.

BPR provides that an applicant may propose to adapt dossier data requirements in accordance with Annex IV, which lists similar options to the REACH Annex XI adaptations, including:

> **Weight of evidence being obtained elsewhere leading to an assumption/conclusion that a substance has or does not have a particular dangerous property. Where such conclusions can be drawn, further testing on vertebrates for that property will not be undertaken**.[118]

That the data should have been 'obtained elsewhere' limits information to that already produced.

Annex III, concerns dossier information requirements for biocidal products (i.e. whole products). Much like Annex II, the Annex reaffirms the importance of reducing vertebrate testing and provides specific indications for the adaption of 'some' data elements. It is notable, that some data elements can be satisfied by 'available information of the properties of active substance(s) contained in the product, and the properties of non -active substance(s) included in the product'. Again, like Annex II, pre-consultation should take place, and new tests involving vertebrates should only be conducted as a last resort. An applicant may propose adaptions in accordance with Annex IV (as outlined above).

Although the BPR signals a reduction in animal testing through the provision of alternatives (in some cases), there remains an inherent bias towards them. This is particularly the case, given that the Regulations provide for the adaption of some, but not all, data points.

5.5 COSMETICS

The Cosmetics Regulation requires a cosmetic product, when being made available on the market, to be safe for human health when used under normal or reasonably foreseeable conditions of use.[119] In undertaking a product safety assessment, the responsible person is required to ensure that the intended use of the cosmetic product, and its anticipated systemic exposure to individual ingredients in the final formulation, are taken account of and that when reviewing data, an appropriate weight of evidence approach is used.[120]

[118] BPR Article 6(3).
[119] Cosmetics Regulation Article 3.
[120] Cosmetics Regulation Article 10, 1 (a) – (b).

Per Annex I, 'relevant' toxicological endpoints should be considered, with a particular focus on local toxicity evaluation (skin and eye irritation), skin sensitisation, and 'in the case of UV absorption, photo-induced toxicity'. All significant toxicological routes of absorption should be considered, as well as a calculation of 'systemic effects and margin of safety (MoS) based on a no observed adverse effects level (NOAEL)'. Reference to 'relevant' toxicological endpoints is notable, given that it implies discretion. Indeed, the European Commission's implementing decision, notes that:

> **The toxicological profile may address a number of different endpoints. A final decision about which endpoints are relevant is made by the safety assessor on a case-by-case basis, taking into account exposure, use of the product, the physico-chemical characteristics of the substances, experience with the substances, etc. Attention should also be paid to local effects (e.g. irritation and photo-toxicity), when relevant. Where a certain endpoint is considered to be not relevant, this should be justified.**[121]

Though toxicological testing is required, Article 18 of the Regulation introduces, 'without prejudice to the general [safety] obligations', a prohibition on the placing on the market of cosmetic products where the final formulation, ingredients in that final formulation or a finished product has been tested on animals using a method other than a validated and adopted alternative method. Notably, however, there may be a requirement for substances, which form part of the cosmetic product, to undergo testing to comply with REACH's information requirement, where no such prohibition on animal testing exists. Given this juxtaposition, ECHA offers the following clarification:

- **Registrants of substances that are exclusively used in cosmetics may not perform animal testing to meet the information requirements of the REACH human health endpoints, with the exception of tests that are done to assess the risks to workers exposed to the substance. Workers in this context, refers to those involved in the production or handling of chemicals on an industrial site, not professional users using cosmetic products as part of their business (e.g. hairdressers).**

[121] Commission Implementing Decision on Guidelines on Annex I to Regulation (EC) No 1223/2009 on Cosmetic Products [2013] OJ L 315/82, Annex 3.8.2.

Registrants of substances that are used for a number of purposes, and not solely in cosmetics, are permitted to perform animal testing, as a last resort, for all human health endpoints.

- **Registrants are permitted to perform animal testing, as a last resort, for all environmental endpoints.**

  Therefore, the testing and marketing bans in the Cosmetics Regulation do not apply to testing required for environmental endpoints, exposure of workers and non-cosmetic uses of substances under REACH.

  Registrants of substances registered exclusively for cosmetic use will still have to provide the required information under REACH wherever possible, by using alternatives to animal testing (such as computer modelling, read-across, weight of evidence etc.).[122]

Further, the Cosmetics Regulation allows that in 'exceptional circumstances' a Member State may request a derogation for the animal testing ban noted above.[123]

Given the above, NAMs, where validated, are accepted for the purposes of the safety report. However, though the Regulation suggests an outright ban on animal testing, subsequent guidance suggests that this is only the case where substances are exclusively used in cosmetics (though note worker exclusions). Where substances offer dual purposes, animal testing is not prohibited, though endpoints are determined on a 'relevant' basis. Although the attempt by the Cosmetics Regulation to ban animal testing can be seen as the low hanging fruit, given that products are not intended to be harmful, this is marred by the interpretation that at the point of production or release to the environment, alternative methods cannot be trusted. This does, however, provide an example of the importance of soft law guidance for industry and the scope for innovation provided.[124]

[122] ECHA, 'Clarity on Interface Between REACH and the Cosmetics Regulation' (*European Chemicals Agency,* 27 October 2014) available at <https://echa.europa.eu/-/clarity-on-interface-between-reach-and-the-cosmetics-regulation> accessed 23 January 2025.
[123] Cosmetics Regulation Article 18.
[124] European Commission, 'SCCS Notes of Guidance for the Testing of Cosmetic Ingredients and their Safety Evaluation – 12th Revision' (*European Union,* 22 December 2023) available at <https://health.ec.europa.eu/publications/sccs-notes-guidance-testing-cosmetic-ingredients-and-their-safety-evaluation-12th-revision_en> accessed 24 January 2025.

5.6 FOOD CONTACT MATERIALS (FCM)

Consisting of a framework regulation and other more specific measures, FCM regulations offer a complex and often disjointed approach to managing chemical risk for materials and articles that will, or are reasonably expected to, come into contact with food. Article 3(1)(a) of the framework regulation (1935/2004) establishes a general requirement for materials and articles to be manufactured in compliance with 'good manufacturing practice' so that they do not endanger human health by transfer of their constituents into food.

Article 5 of the Regulation provides for the adoption of 'specific measures' for groups of materials and articles listed in Annex 1; those being active and intelligent materials and articles, adhesives, ceramics, cork, rubbers, glass, ion-exchange resins, metals and alloys, paper and board, plastics, printing inks, regenerated cellulose, silicones, textiles, varnishes and coatings, waxes, wood. Where adopted, these specific measures may include a list of substances authorised for use in the manufacturing of materials and articles and specific limits on the migration of certain constituents into or onto food.[125]

Article 9 requires an application for approval to include a technical dossier which contains 'information specified in the guidelines for the safety assessment of a substance to be published by' the EFSA, and this guidance is available for adopted specific measures. Taking the examples of specific measures on plastics (10/2011) and active and intelligent materials and articles (450/2009), respective guidance (which is supported by the force of Article 9 of the framework regulation) requires a risk assessment, including exposure (see Annex III of this document).

Neither the framework regulation, nor the example specific measures analysed in this document, explicitly provide for the reduction of animal testing nor do they actively encourage the use of NAMs. However, taking plastics as an example, associated guidance operates on the premise of the 'greater the exposure to the substance through migration, the more toxicological information will be needed' and thus, testing requirements are dependent on migration.[126] In case of high migration (5 - 60 mg/kg/food), a full data set is needed, though it 'may' be sufficient for this to be reduced, in cases of migration between 0.05 and 5 mg/kg food. For low migration (<0.05 mg/kg food), the guidance provides that 'only a limited data set is needed'. In cases of high and reduced migration, animal testing is required, with, for example, both categories stipulating a 90-day oral toxicity study. With low migration, animal testing is not guaranteed, though may in some cases

[125] FCM Article 5 (1)(a) and (e).
[126] EFSA Panel, 'Note for Guidance for the Preparation of an Application for the Safety Assessment of a Substance to be used in Plastic Food Contact Materials' (2008) 6(7) EFSA Journal 41, 6.

be required, where a substance undergoing genotoxicity testing returns 'positive in the in vitro basic battery'.

Therefore, while REACH is typically central to the NAM debate, most likely due to its greater use by industry, it appears that the product-specific legislation requires considerable rethinking to ensure alternatives are used. Having said that, the legislation does provide a platform for further soft law guidance which might be employed to widen the scope for new test methods.

5.7 PHARMACEUTICALS

Pharmaceuticals are comprised of biologicals or biotherapeutics (such as vaccines), and medicines. Risk appetites may vary, between medicines taken to treat illness and disease in an attempt to make a person better, and vaccines, which are made for healthy people to prevent them from becoming ill. That being said, both are subject to authorisation, which requires information including adverse reactions and toxicological tests.[127]

Initial research on medicines is typically undertaken by pharmaceutical and biotechnology companies, first tested in a laboratory (including non-clinical trials) and then in human volunteers or patients (clinical trials).[128] Applications to conduct clinical trials are directed to Member States rather than the European Medicines Agency (EMA), which is the body that provides fee-payable scientific advice on study design (complementary to general guidelines) and also makes the resulting decision on marketing authorisation.[129]

Testing may be in vitro or in vivo and they may cover repeated dose toxicity (sub-acute, sub-chronic toxicity, or chronic toxicity), carcinogenicity, developmental toxicity, and reproductive toxicity.[130] Information is used, for example, to estimate an initial safe dose for human trials and for monitoring potential adverse effects.[131] The principles of the 3Rs as per Directive 2010/63 are

[127] European Union Directive 2001/83/EC on the Community Code Relating to Medicinal Products for Human Use [2001] OJ L 311/67 ('2001/83'), Article 8.
[128] EMA, 'From Laboratory to Patient: The Journey of a Medicine Assessed by EMA' (*European Medicines Agency,* 2019) Available at <https://www.ema.europa.eu/en/documents/other/laboratory-patient-journey-centrally-authorised-medicine_en.pdf> accessed 22 January 2025; Regulation (EU) No 536/2014 on Clinical Trials on Medicinal Products for Human Use [2014] OJ L 158/1.
[129] Ibid, EMA. See also EMA, 'Scientific Advice and Protocol Assistance' (*European Union,* 10 January 2025) available at <https://www.ema.europa.eu/en/human-regulatory-overview/research-development/scientific-advice-protocol-assistance> accessed 27 January 2025.
[130] Doortje Swaters, and others, 'A History of Regulatory Testing: What Can We Learn?' (2022) 50(5) Alternatives to Laboratory Animals 322.
[131] Ibid. See also EMA, 'Guidelines on Strategies to Identify and Mitigate Risks for First In-human and Early Clinical Trials with Investigational Medicinal Products' (*European Union,* 20 July 2017) available at <https://www.ema.europa.eu/en/documents/scientific-guideline/guideline-strategies-identify-and-mitigate-

also applicable to pharmaceuticals and is expected to be upheld in European Pharmacopoeia monographs (Ph. Eur.), which are considered the regulatory accepted guidelines along with other guidance adopted by the Committee for Medicinal Products for Human Use (CHMP) of the EMA.[132] Scientific validity may be demonstrated through validation by ECVAM or following a case-by-case evaluation.[133] Relevance of alternative methods for non-clinical tests takes a context-of-use approach, where a 'description of circumstances' in which the method is applicable is required;[134] a qualification approach involving a public consultation to gather the views of the scientific community of non-test-guideline methods is also possible.[135]

The International Council for Harmonisation of Technical Requirements for Pharmaceuticals for Human Use (ICH) provides internationally harmonised guidelines for quality, safety, and efficacy testing of human medicinal products, including principles such as for decision making to be science-based. The EMA and the US Food and Drug Administration (FDA), amongst other regions, are part of ICH and WHO.[136] The European Directorate for the Quality of Medicines (EDQM) is responsible for the Ph. Eur, which is considered legally binding.[137]

Tests may deviate from EMA or ICH guidelines on the basis of scientific justification. If this concerns the use of data provided by a 3Rs approach, the test are to provide 'an equivalent level of quality, safety, or efficacy,' or a justification that a test is not necessary.[138] The most recent overview of current regulatory testing requirements and opportunities for implementation of the 3RS is from 2018: under the Safety Working Party of the CHMP, requirements do include reference to species in repeated dose toxicity testing and rats and mice in carcinogenicity

---

risks-first-human-and-early-clinical-trials-investigational-medicinal-products-revision-1_en.pdf> accessed 27 January 2025.

[132] EMA, 'Guideline on the principles of regulatory acceptance of 3Rs (replacement, reduction, refinement) testing approaches' (*European Medicines Agency,* 15 December 2016) Available at <https://www.ema.europa.eu/en/documents/scientific-guideline/guideline-principles-regulatory-acceptance-3rs-replacement-reduction-refinement-testing-approaches_en.pdf> accessed 22 January 2025.

[133] Ibid.

[134] EMA, n132, p7.

[135] Ibid, section 5.5.1.

[136] However, at the time of writing, President Donald Trump has issued an Executive Order that the United States intends to withdraw from the WHO: 'Withdrawing the United States from the World Health Organisation' (*The White House*, 20 January 2025) available at https://www.whitehouse.gov/presidential-actions/2025/01/withdrawing-the-united-states-from-the-worldhealth-organization/> accessed 24 January 2025.

[137] 2001/83, n127.

[138] EMA, 'Reflection Paper Providing an Overview of the Current Regulatory Testing Requirements for Medicinal Products for Human Use and Opportunities for Implementation of the 3Rs' (*European Medicines Agency,* 18 October 2018) available at <https://www.ema.europa.eu/en/documents/scientific-guideline/reflection-paper-providing-overview-current-regulatory-testing-requirements-medicinal-products-human-use-and-opportunities-implementation-3rs-first_en.pdf> accessed 23 January 2025.

testing.[139] Currently, in some specific examples some in vitro testing is recommended rather than animal testing, such as the battery for genotoxicity testing; one stage of development within reproductive toxicity; and QT prolongation (irregular heart rhythm).

In terms of vaccines, the Ph. Eur 5.2.14 monograph describes what is sought for appropriately designed in vitro substitution in the quality controls (QC) of vaccines. This is mostly a problem for legacy vaccines, rather than new vaccines and biotherapeutics, whose resort to in vivo is rare. General Notices relay information explaining which statements are mandatory (or not), providing definitions of terms. They also support alternatives, and methods appearing in General Notices are considered validated (hence some inclusion of in vivo). It is acknowledged in 5.2.14 that in vitro assessment is different from in vivo, e.g., rather than potency the in vitro will look at antigen content or a functional response such as a viral neutralisation; and rather than in vivo toxicity the in vitro will show toxin binding and enzyme activation. In vitro is expected to be validated, and a battery of methods is acceptable to show characteristics. For toxicity, in vitro is to be specific and as sensitive as in vivo and be a functional system such as a toxin-sensitive cell line or be linked to mode of action (e.g., receptor binding and enzyme activity).

As with chemicals, most in vivo tests pre-date guidelines on validation (ICH Q2(R1) or VICH GL2) and they are instead considered validated because they are part of a compendium of methods.[140] There also appear to be deep-seated beliefs: for example, despite not being scientifically justified, there is considered to be a 'belief' held by some regulators that in vivo is central to maintaining the safety of vaccines.[141] Despite this, the Pertussis Histamine Sensitisation Test and some other in vivo tests have now been removed from Ph. Eur (also Tetanus specific toxicity, Diphtheria specific toxicity, etc.). Non in vivo QC tests include the HPV vaccine (NIH animal assay is unreliable), meningococcal and Pneumococcal bacterial conjugate vaccines (uses a combination of methods), and EMA and North American Covid-19 vaccines.[142]

Therefore, while steps are being taken to avoid and replace animal tests in medicine development, some remains. While the Ph. Eur. provides a focal point for accepted methods for quality tests, case-by-case assessments are possible. The use of alternative methods for new

[139] Ibid. This is now the Non-clinical Working Party, and the reflections documents are being updated, with public consultation on them due soon.
[140] Dean Smith, 'Substituting In Vitro for In Vivo Potency and Safety Assays: Sciences Versus the Fear Factor' (The Humane Society of the United States: Transition to Non-animal-based Vaccine Batch Release Testing, webinar, March 2024) available at <https://www.afsacollaboration.org/sciencex_event/transition-to-non-animal-based-vaccine-batch-release-testing/> accessed 23 January 2025.
[141] Ibid.
[142] Ibid.

vaccines appears more accepted, and the alternative routes possible of considering a method to be acceptable (such as by qualification) are encouraging.

5.8 CHAPTER CONCLUSION

While each legislative instrument expects a high level of protection and states that the use of animals is primarily as a last resort, the testing requirements lean very heavily on traditional animal testing. Indeed, most legislation make strong references to specific animal studies. There does, however, appear to be greater scope for the use of NAMs under REACH, thanks to the Annex XI adaptations and WoE in particular. This is likewise the case for CLP, which defies general opinion. Instead, the greater issue here is not the legal text, but its interpretation. There is less scope to utilise NAMs under BPR, and PPP diverges to the greatest extent from the 2010/63 Directive through its more direct reference to animal studies with little ability to avoid animal tests. Therefore, these and the FCM, which may be considered lesser-mentioned regulations in contrast to REACH and CLP, appear to have the furthest to go, in terms of legal amendment to accommodating alternatives. However, both these and the Cosmetics Regulations do exemplify the greater role that soft law guidance can play in supporting regulations and providing authority to utilise NAMs.

A point worth noting is with regards to the frequent mentions of validated methods. Validation is not legally defined with the legislation, although expectations can be informed by descriptive text, such as that from REACH: '[methods in TMR] or in accordance with other international test methods recognised by the Commission or [ECHA] as being appropriate.' Likewise, in CLP tests are to be conducted either as those acceptable under REACH or following 'sound scientific principles that are internationally recognised or methods validated according to international procedures.' Current interpretation is that OECD GD34 defines validation, and validity stems from OECD members' consensus of the text of this guideline,[143] yet other routes to acceptance could be developed or other existing principles utilised, with inspiration taken from the pharmaceutical sector. Regulators may state their preference to be OECD TGs due to international consensus and the mutual acceptance of data (MAD), claiming that TGs alone provide an expedited process may be overstated. However, prioritisation may prove a more efficient route rather than an exclusive focus on TGs, as the substance evaluation process is not

[143] OECD, 'Guidance Document on the Validation and International Acceptance of New or Updated Test Methods for Hazard Assessment' (*OECD,* 18 August 2005) available at <https://www.oecd.org/en/publications/guidance-document-on-the-validation-and-international-acceptance-of-new-or-updated-test-methods-for-hazard-assessment_e1f1244b-en.html> accessed 23 January 2025.

swift and is already undertaken on a case-by-case basis. While standardisation provides this to an extent, validation need not necessarily be the sole route.[144] Additionally, while harmonisation may be desirable to support MAD, the resulting decisions based on mutually accepted data continue to vary between jurisdictions. It may be the case, then, that international effort might prove more effective in focussing on reducing animal use by avoiding repetitive testing,[145] rather than chasing harmonisation to support market access.[146] Ultimately, this legislative mapping has demonstrated that NAMs can be accommodated within key pieces of regulatory frameworks, and recommendations on how to improve the uptake of NAMs more widely can be identified.[147]

6. Conclusion

This report has considered the capacity of regulatory frameworks to accommodate NAMs through first considering the state-of-play of animal testing under the relevant legislation, from which it is clear thousands of animals continue to be used annually. The risk assessment critique identified the value that can be added from NAMs, for example providing depth of knowledge from mechanistic information, which is not fully recognised by EU law as it is currently being implemented. Further to this, the case law analysis questioned the extent to which the last resort principle was applied in oversight of regulatory decision making. Finally, the legislative mapping identified some, if limited, scope to implement NAMs across chemical legislation.

While there are many benefits of OECD TGs for regulators, there is a strong case for policymakers to improve alignment with the 2010/63 Directive and legislated aims of reducing animal testing in regulations and ensuring protection. While regulators may wish to see a link between activity and an adverse effect, this is an interpretation rather than a requirement mandated in legislation, and one that, based on the challenges to NAM use identified in the case law analysis, up to now

[144] Joint Research Centre, 'Good Practices and Resources to Improve the Utility of Research Data in Regulatory Assessments: Webinar Outcome Report' (*European Commission,* 2024) available at <https://joint-research-centre.ec.europa.eu/document/download/b11c36cc-7611-4a4e-aa78-926bb64edc43_en?filename=JRC137088%20-%20Good_practices_research_data_regulatory_assessments.pdf> accessed 23 January 2025.

[145] Although this is also not guaranteed: see ECHA Board of Appeal Digest at n104. Para 8.4: "OECD decision on the mutual acceptance of data (MAD): … The MAD system is not binding on the Agency as the European Union has not acceded to the Convention on the OECD".

[146] States take action at different rates, for example restrictions on chrysotile asbestos: EPA Chemical Update, 'Biden-Harris Administration Finalizes Ban on Ongoing Uses of Asbestos to Protect People from Cancer' (*EPA's Office of Chemical Safety and Pollution,* 18 March 2024) see <https://www.epa.gov/assessing-and-managing-chemicals-under-tsca/risk-management-asbestos-part-1-chrysotile-asbestos> accessed 23 January 2025.

[147] While the doctrinal analysis was conducted by legal scholars, advice on the text produced for this report was sought from relevant actors, to ensure accuracy. Additionally, a detailed mapping of REACH can be found at Annex IV.

appears opaque. Without clearer policy direction, some of which may be provided by utilising soft law guidance rather than hard law legislative change, there is a risk that industry will continue to feel uncertain about the inclusion of NAM data in dossiers, which inherently stunts the advancement of new methods.

This report is timely, coming as the European Commission develops its roadmap to phase out animal testing, and we can suggest two areas of further research: firstly, as mentioned in Chapter 2 above, a qualitative empirical study to discern reasons for variable animal use data; and secondly, research on alternative internationally recognised principles on acceptable methods, either from the OECD or other institutions such as ISO. We would also recommend the EC draws lessons from approaches to replace animal testing as pursued in the Cosmetics Regulation and pharmaceutical sectors.

Annex I: Tables of Case Law

With contributions from Rubina Sultan-Chaudhary

Annex II: Overview of Key Attributes of Mapped Legislation

With contributions from Louis Dawson

Annex III: Comparison of Toxicological and Ecotoxicological Endpoints by Legislation

With contributions from Louis Dawson

Annex IV: Legislative Mapping of REACH

List of acronyms

Abbreviation	Description
3Rs	Refine, Reduce, Replace animal testing
AOEL	Acceptable Operator Exposure Level
AOP	Adverse Outcome Pathway
ALURES	Animal Use Reporting – EU System
ANSES	Agence Nationale de Sécurité Sanitaire de l'alimentation, de l'environnement et du Travail
BPR	Biocidal Products Regulation
CHMP	Committee for Medicinal Products for Human Use

CJEU	Court of Justice of the European Union
CLP	Classification, Labelling, and Packaging Regulation
CMR	Carcinogens, Mutagens, and Reproductive toxicants
DA	Defined Approaches
EC	European Commission
ECEAE	European Coalition to End Animal Experiments
ECHA	European Chemicals Agency
ECVAM	EU Reference Laboratory for Alternatives to Animal Testing (EURL ECVAM)
EDQM	European Directorate for the Quality of Medicines
EFSA	European Food Safety Authority
EMA	European Medicines Agency
EU	European Union
FCM	Food Contact Materials Regulation
FDA	Food and Drug Administration of the United States
GHS	Globally Harmonized System of the United Nations
GFRB	Group First, Regulate Better
IATA	Integrated Approaches to Testing and Assessment
ICH	International Council for Harmonisation of Technical Requirements for Pharmaceuticals for Human Use
ISO	International Organisation for Standardisation
JRC	Joint Research Centre of the European Commission
MAD	Mutual Acceptance of Data
MKE / KE	Molecular Key Event / Key Event
MoS	Margin of Safety
MSC	Member State Committee
NOAEL	No Observed Adverse Effects Level
NAMs	New Approach Methodology
NGO	Non-Governmental Organisation(s)
NGRA	Next-Generation Risk Assessment
OECD TG	Organisation for Economic Co-operation and Development Test Guideline(s)
PARC	Partnership for the Assessment of Risks from Chemicals

PFAS	Perfluoroalkyl and Polyfluoroalkyl Substances
PFOS	Perfluorooctane Sulfonic Acid
PoD	Point of Departure
POPs	Persistent Organic Pollutants Regulation
PPP	Plant Protection Product Regulation
QC	Quality Controls
RAC	Committee for Risk Assessment
REACH	Registration, Evaluation, Authorisation and Restriction of Chemicals Regulation
RIVM	National Institute for Public Health and the Environment, Netherlands
TFEU	Treaty for the Functioning of the European Union
TMR	Test Method Regulation
USA	United States of America
WHO	World Health Organisation
WoE	Weight of Evidence