All information regarding future IHI Call topics is indicative and subject to change. Final information about future IHI Calls will be communicated after approval by the IHI Governing Board.

Topic 2 : An AI Foundation Toxicology Model and Framework to Support Waiving a Second Species in Drug Safety Studies

Expected outcomes

The action under this topic must contribute to all of the following outcomes:

1. **A validated Artificial Intelligence (AI) Foundation Toxicology Model** that provides transparent probabilistic predictions for industry and regulator stakeholders to determine when a second species in chronic (>90 days) and sub-chronic (90 days) small molecule medicine repeat-dose studies is unlikely to provide additional safety relevant information, including risks of missed toxicity, organ-specific findings, and divergence in No Observed Adverse Effect Level (NOAEL). The goal would be to enable waiving the need for two-species chronic testing for small molecules and other modalities e.g. oligonucleotides.

2. **A standardised, transparent weight-of-evidence framework** for industry, regulator and academic stakeholders that enables reproducible assessment of evidence quality, consistency, relevance, and uncertainty across regulatory submissions, supporting the wider adoption of the AI Foundation Toxicology Model and New Approach Methodology (NAM)-based toxicology strategies in general.

3. **Functional tools, templates, and training materials** that support the real-world implementation, sustainability and evolution of the foundation model and weight-of-evidence framework including guidance on explainability, provenance, governance, ethical use, and alignment with AI requirements, tailored to industry, regulator and academic stakeholder needs.

4. **Enhanced industry and regulator stakeholder confidence in second species waiver applications**, particularly for small molecule medicines, supported by empirical, calibrated evidence and a framework enabling predictable adjudication, more consistent global waiver decision-making and timely progression of medicine development without compromising patient safety. This confidence should be gained through the model and framework's application for regulator validation and acceptance, with the longer-term goal of revising the regulatory guidelines ICH M3(R2) **(1)** to align with this project's outcomes. This topic should provide the opportunity to extend this confidence to waiving chronic testing in a single species beyond small molecule medicines and to other study types.

Scope

Regulatory guidelines e.g. ICHM3(R2) for nonclinical safety assessment of new small molecule medicines have historically required toxicity studies in two species, typically one rodent and one non rodent, to maximise the likelihood of detecting adverse effects, improve the translatability to humans and guide patient dosing and monitoring. This two species approach is increasingly challenged by ethical considerations, long and costly study durations, limited reproducibility, and uncertainty regarding human relevance **(2,3)**. In parallel, scientific progress in mechanistic toxicology, advanced *in vitro* systems (e.g. organoids and micro-physiological systems), multi-omics technologies, and AI-based modelling is enabling NAMs to help reduce reliance on animal models without compromising patient safety. Within the context of this topic, the AI Foundation Toxicology Model will be considered a NAM, facilitating NAM-enabled second species waivers.

In addition, the policy landscape across global jurisdictions is evolving rapidly. The European Commission is preparing a Roadmap towards phasing out animal testing for chemical assessment safety, expected in 2026 **(4)**; the EMA has increased its focus on emerging NAMs through horizon-scanning and dedicated working groups **(5)**; the UK Government has published the Replacing Animals in Science strategy **(6)** and the FDA has published a dedicated roadmap for reducing animal testing in preclinical safety studies **(7)**. Collectively, these initiatives highlight the need for robust evidence, data integration, validated methods and coordinated stakeholder action to support credible alternatives to animal testing.

A weight-of-evidence approach can eliminate unnecessary animal tests by integrating and assessing diverse data sets as well as determining whether the available evidence is sufficient or additional nonclinical testing is required to address potential safety concerns. The value of this evidence depends on its quality, consistency, human relevance, and the nature and severity of observed effects. Currently, the absence of standardised, data-driven methods means decisions can vary across reviewers and jurisdictions, creating an uncertainty which discourages changes from being made to the current process. A more structured, reproducible, validated, and evidence-based approach would support consistent, transparent, and well justified decision-making on whether second species studies could be waived.

This topic aims to address and standardise the waiver process for second species chronic and sub-chronic small molecule general toxicology studies by developing a transparent, AI-driven Foundation Toxicology Model. The model should be underpinned by diverse, standardised, high quality *in vivo* datasets, encompassing all modalities of medicines, and supported by a robust weight-of-evidence framework. The most appropriate single species should be selected early in the drug development programme based on target affinity and functions, metabolite profiling, etc. The main context of use of the foundation model is to predict whether testing in a second species will provide additional safety relevant information. If toxicity predictions are equal across rodent and nonrodent species, priority should be given to the reduction of nonrodent chronic toxicity testing.

Although the context of use for this topic will be focused on waiving second species testing for chronic and sub-chronic small molecule studies, the project's actions, sustainability plan plus the ongoing development and evolution of the model should provide the opportunity to assess appropriate single species selection and encompass the broadening of the context of use to enable the waiving of single species chronic testing in small molecules and other modalities, such as oligonucleotides. Therefore, relevant data sources and their provision should enable this broader and longer-term goal. Selection of the appropriate single animal species for preclinical safety evaluations for biotechnology-derived pharmaceuticals in line with ICH S6 **(8)** should be considered within the scope of this project if the relevant data provision is sufficient to enable this functionality.

This topic will build upon, among others, the outcomes, learnings, data sharing structures, and methodologies of prior IMI/IHI projects that improve the predictability, feasibility and reliability of pre-clinical safety assessments—such as BigPicture **(9)**, eTRANSAFE **(10)**, eTOX **(11)**, imSAVAR **(12)**, and VICT3R (**13**)—as well as other relevant European, international and national initiatives such as Animal-free safety assessment of chemicals: Project cluster for implementation of novel strategies (A SPIS) **(14)** and NC3Rs-led efforts like the Virtual Second Species **(15)** initiative and the Two Species project **(16)**. Preliminary analysis from the NC3Rs Two Species project predicts that it would be possible to avoid conducting studies in non-rodents (e.g. dog or non-human primate) in >60% of the small molecule drugs that enter Phase II globally each year **(17)**.

By fostering collaboration among pharmaceutical companies, regulatory agencies, academic institutions, SMEs, and patient advocacy groups, the action generated by this topic should seek to enhance patient safety, reduce reliance on animal testing, streamline medicine development timelines, lower associated costs, and boost the global competitiveness of the pharmaceutical industry within the European Union.

To fulfil this aim, the action should:

1. **Consolidate and evaluate data sources**
    - Identify, curate, and assess the quality and suitability of data from multiple pharmaceutical companies and other organisations within the consortium with a priority focus on general toxicology studies (14-day, 28-day, 90-day and chronic toxicity studies) in the form of structured toxicity data in CDISC SEND format as well as in unstructured study reports. (SEND specifies a way to collect and present nonclinical data in a consistent format, a requirement for data submission to regulators)**.** Secondary focus for additional data should also be considered including *in vivo* toxicity study data, chemistry (e.g. structure information, where permitted), pharmacokinetic and exposure data, mechanistic *in vitro* systems, *in silico* models, multiomics datasets, licensed content, prior consortia outputs, and public regulatory databases.
    - Develop a new database to host and analyse data, encompassing principles to support a common or federated data model that enables sensitive data preservation and multisource analysis that can allow NAM-enabled second species waivers. Additionally, sensitive data should undergo various levels of blinding from full blinding to structural alerts as a surrogate for the full structure and structural embeddings with noise introduction among other approaches.

- Applicants are expected to consider the potential regulatory impact of the project's results and develop a regulatory strategy and interaction plan for generating the required evidence to support regulatory decision-making, linked to the data sources used, as well as engaging with regulators in a timely manner for their input (e.g. national competent authorities, EMA, etc). The goal is to gain, by the end of the project, endorsement of the AI Foundation Toxicology Model from the regulatory authorities (e.g. EMA, FDA, including potential ISTAND submission) to enable second species waivers for chronic and sub-chronic small molecule testing. The longer-term goal of revising the regulatory guidance ICH M3(R2) in line with this project's outcomes should also be part of the regulatory strategy, noting that whilst this is unlikely to happen during the project's duration, the foundations should be built. The opportunity to broaden the scope of the Foundation Toxicology Model to encompass waiving single-species chronic testing beyond small molecules e.g. oligonucleotides should be part of the longer-term and sustainability planning.

2. Select and optimise AI methods

- Conduct a systematic review of existing AI and AI-supported foundation modelling approaches.
- Select and optimise AI methods able to perform probabilistic predictions estimating the likelihood of novel safety-relevant information being identified in longer-term studies and/or in a second species, including risks of missed toxicity, organ-specific findings, and divergence in NOAEL. The potential to select the appropriate single species for biological medicines should also be considered in the AI method selection and optimisation.
- Ensure transparency, interpretability, traceability, and data provenance to meet regulatory expectations.

Applicants should ensure that, in order to support regulatory acceptance and the scientific robustness of the Foundation Toxicology Model, it is built using a structured and transparent AI methodology encompassing:

- Data Harmonisation and Representation Learning
  - Feature extraction pipelines appropriate to each data modality (unstructured study reports, molecular structures, omics datasets, mechanistic assays).
  - Multimodal representation learning (e.g., contrastive learning) to create unified latent spaces across biological, toxicological, and metadata features.
  - Federated learning and privacy preserving techniques to enable secure, multi-organisation data contributions without exposing proprietary information.
- Predictive Modelling Approaches
  - Probabilistic supervised models (e.g., Bayesian neural networks, Gaussian Processes, calibrated ensembles) to estimate the probability of missed toxicological findings under a single-species approach.
  - Mechanistic-informed modules integrated via knowledge graph neural networks, causal inference frameworks, and multitask predictors for endpoints such as drug-induced liver injury, cardiotoxicity, and genotoxicity.

    - Uncertainty quantification methods to ensure transparent confidence intervals suitable for risk-based decision-making.
    - Scenario simulation engines (Monte Carlo, generative models, causal models) to test counterfactuals relevant to second species value.
- Transparency and Explainability
    - Use of explainable AI techniques to identify key drivers of model predictions.
    - Comprehensive provenance tracking to ensure every model output is traceable.
    - Human interpretable decision layers to support weight-of-evidence narratives.
- Long-Term Evolution
    - Ability for continuous learning, single-species chronic testing waivers beyond small molecules and modular updates as new data types and methods emerge.

In parallel, other approaches should be assessed in case a foundation model does not provide the solution. This could encompass decision-tree approaches, Bayesian statistical methods and/or more classical Machine Learning (ML) approaches.

3. **Develop and validate the foundation toxicology model**

- Design, train, and validate a foundation toxicology model capable of providing probabilistic predictions across a number of tasks that span study duration (sub-acute (14-28 days) to sub-chronic to chronic), species (rodent vs non-rodent) and prediction output including the likelihood of identifying a novel target organ, a significantly divergent toxicity profile, and/or a significantly divergent NOAEL. For example, the model should predict the likelihood (0-100) that running a 9-month non-rodent study would identify a divergent toxicity profile, including significantly increased severity of pathology findings as compared to the toxicity profile identified in earlier, shorter-term studies or projected to be identified in the rodent 6-month study yet to be conducted. These specific outputs will be defined depending on the data provided and the AI methods, which will be established within the early stages of the project.
- Industry, academic and SME stakeholders with consistent regulator engagement and input should define representative use cases and performance metrics, including accuracy, robustness, explainability, extensibility and trustworthiness, to validate the model. Formal benchmarking and rigorous testing are essential for ensuring consistency of results, to build trust in the model's recommendations.

4. **Establish a standardised weight-of-evidence framework**

- Create a transparent, reproducible, weight-of-evidence decision support framework that integrates diverse data types and model outputs into standardised reasoning steps, including the direct model outputs (e.g., probability and significance of novel findings in a longer-term second species study) and model context (e.g., narrative rationale supporting the probabilistic output) from the foundation model as well as orthogonal internal programme and external literature support. *In vitro* NAMs, target knowledge and literature mining would all help supplement and increase confidence in the Foundation Toxicology Model outputs and recommendations.
- Define how uncertainty, evidence quality, consistency, severity, and human relevance including potential impact on or risk for the clinical programme should be evaluated, compatible with regulatory reasoning, to support second species waiver submissions.
- Test the foundation model, through use cases, together with the weight of evidence framework with relevant diverse stakeholder groups, including pharmaceutical and biotechnology companies, regulators and academics/SMEs working at the interface of drug safety research and regulation to ensure consistency across reviewers and jurisdictions.

5. **Prepare for regulatory uptake and long-term sustainability**

- Develop recommendations and practical tools for real-world implementation, including regulatory strategy and guideline revision as required, alignment with ethical and legal principles plus a governance structure for ongoing model evolution and ecosystem adoption. For instance, trustworthy AI, human oversight and verifications will follow regulatory frameworks such as the Assessment List for Trustworthy Artificial Intelligence (ALTAI)**(18)**.

- Devise a sustainability and evolution plan for long-term hosting, maintenance, continuous data harmonisation and improvement of the foundation model and weight-of-evidence framework, enabling broad and sustainable ecosystem adoption.

Expected impacts

The action under this topic is expected to achieve the following impacts:

- Faster and more informed decision-making through the use of an AI-driven NAM (Foundation Toxicology Model) and increased efficiency through rapid processing of vast amounts of data.
- Increased consistency and standardisation in a NAM-based approach, specifically an AI model, used by industry in the efficient development, testing and production of safe and effective innovative health technologies, improving industrial competitiveness.
- Regulatory adoption of a NAM-enabled second species waiver model (AI Foundation Toxicology Model) and weight-of-evidence framework, in line with recommendations and more consistent global decision-making on waiving second species testing.
- Reduction in animal use, accelerated timelines and lower costs, enhancing the competitiveness of the European health industry through economical and ethical benefits.
- Improved public health as patients will benefit from safe and effective medicines developed faster using validated NAMs.

The action is expected to contribute to the EU Directive (2010/63/EU)(**19**) on the protection of animals used for scientific purposes and the implementation of the 3Rs principles to replace, reduce and refine the use of animals. The action is expected also to consider and contribute to EU programmes, initiatives and policies on New Approach Methodologies (NAMs) such as the future European Research Area (ERA) action on accelerating NAMs to advance biomedical research and testing of medicinal products and medical devices

Why the expected outcomes can only be achieved by an IHI JU action

Achieving regulator-trusted, NAM-enabled second species waivers requires large, heterogeneous datasets; federated collaboration across multiple pharmaceutical companies and other relevant stakeholders; cross-disciplinary expertise in toxicology, data science, regulatory sciences, and ethics; and early, coordinated engagement with regulatory authorities. These conditions cannot be met by individual organisations acting independently. IHI JU Public-Private partnerships provide the framework, scale, and infrastructure required to deliver impact, ensure sustainability, and build broad stakeholder confidence in AI driven weight-of-evidence approaches.

Pre-Identified industry consortium

The pre-identified industry consortium that will contribute to this cross-sectoral IHI JU project is composed of the following pharmaceutical and medical technology industry beneficiaries ('constituent or affiliated entities of private members').

In the spirit of partnership, and to reflect how IHI JU two-stage call topics are built upon identified scientific priorities agreed together with a number of proposing industry beneficiaries (i.e. beneficiaries who are constituent or affiliated entities of a private member of IHI JU), it is envisaged that IHI JU proposals and actions may allocate a leading role within the consortium to an industry beneficiary. Within an applicant consortium discussing the full proposal to be submitted for stage 2, it is expected that one of the industry beneficiaries may become the project leader. Therefore, to facilitate the formation of the final consortium, all beneficiaries, affiliated entities, and associated partners are encouraged to discuss the weighting of responsibilities and priorities with regard to such leadership roles. Until the role is formalised by execution of the Grant Agreement, one of the proposing industry beneficiaries shall as project leader facilitate an efficient drafting and negotiation of project content and required agreements.

Indicative budget

- The maximum financial contribution from the IHI JU is up to EUR 9 000 000. ***NB: this amount is indicative and subject to change, pending approval by the IHI Governing Board.***
- The indicative in-kind and financial contribution from industry beneficiaries is EUR 7 300 000 (target). ***NB: this amount is indicative and subject to change, pending approval by the IHI Governing Board.***

Due to the global nature of the participating industry partners, it is anticipated that some elements of the contributions will be in-kind contributions to operational activities (IKOP) from those countries that are neither part of the EU nor associated to the Horizon Europe programme.

The indicative in-kind contribution from industry beneficiaries may include in-kind contributions to additional activities (IKAA).

- The indicative in-kind contribution from the IHI JU contributing partner is EUR 100 000. ***NB: this amount is indicative and subject to change, pending approval by the IHI Governing Board.***

Indicative duration of the action

The indicative duration of the action is 60 months.

This duration is indicative only. At the second stage, the consortium selected at the first stage and the predefined industry consortium and contributing partner may jointly agree on a different duration when submitting the full proposal.

Contribution of the pre-identified industry consortium

The pre-identified industry consortium and contributing partner expect to contribute to the IHI JU project by providing the following expertise and assets:

- Proprietary toxicology and mechanistic datasets
- Expertise in CDISC SEND standards, AI/ML, mechanistic toxicology, pharmaceutical and regulatory sciences, and federated analytics
- Infrastructure for federated learning

- Leadership in model development, validation, and regulatory engagement

Applicant consortium

The first stage applicant consortium is expected, in the short proposal, to address the scope and deliver on the expected outcomes of the topic, taking into account the expected contribution from the pre-identified industry consortium and contributing partner.

This may require mobilising the following expertise:

- Pharmaceutical and toxicology data management, engineering and integration including:
    - Managing chemical structure and properties, structured (SEND-formatted) and unstructured (study reports) *in vivo* toxicology study data, *in vitro* assay results and flexibility to handle other data types and formats (e.g., omics)
    - Mapping of diverse terminology to ontologies and controlled vocabularies for enhanced data integration at scale via automated methods (i.e., LLM-driven)
    - Extraction of key values and text from unstructured reports (e.g., treatment-related findings and NOAEL/LOAEL values)
    - Hosting federated data / learning platforms or other data security and privacy platforms/approaches
- AI/ML model development capabilities including:
    - Core machine learning, artificial intelligence, data science applied to health and biological sciences
    - Advanced AI techniques including ability to develop large-scale models (multi-billion parameter networks)
    - Experience with multi-modal model development
- Experience with complex probabilistic predictions as outputs from foundation models or fine-tuned versions
- Regulatory and compliance Framework Standard Operating Procedures (SOPs): expertise for AI model validation and monitoring
- Regulatory packages and weight of evidence decision making in pharmaceutical toxicology studies
- Project management experience for large multi-stakeholder European Public-Private partnerships

Furthermore, the applicant consortium is expected to provide the below resources:

- Technological infrastructure:
    - High-performance computing for processing large volumes of data
    - Data storage solutions to securely store and manage data from various sources during the project duration and beyond to enable sustainability of the project's outcomes.

    - Federated data platform that enables sharing and / or analysis of data in a blinded manner
- Provision of proprietary toxicology and mechanistic datasets or other relevant data, where available

All project members need to have access to data in a GDPR-compliant manner

The applicant consortium is expected to enable effective collaboration with regulatory authorities, national competent authorities, and may consider, for instance, engaging them as consortium partners, or in an advisory capacity.

At the second stage, the consortium selected at the first stage and the predefined industry consortium and contributing partner will form the full consortium. The full consortium will develop the full proposal in partnership, including the overall structure of the work plan and the work packages, based upon the short proposal selected at the first stage.

Dissemination and exploitation obligations

The specific obligations described in the conditions of the calls and call management rules under 'Specific conditions on availability, accessibility and affordability' do not apply.

References

1. ICH M3(R2). Nonclinical safety studies for the conduct of human clinical trials and marketing authorization for pharmaceuticals. In: *International Conference on Harmonisation (ICH)*. Topic M3(R2); 2009.
2. Prior *et al.* (2020). Opportunities for use of one species for longer-term toxicology testing during drug development: a cross industry evaluation. *Regulatory Toxicology and Pharmacology.* 29;113:104624
3. Prior *et al.,* (2024). Exploring Greater Flexibility for Chronic Toxicity Study Designs to Support Human Safety Assessment While Balancing 3Rs Considerations *International Journal of Toxicology.* 43(5):456-463. https://journals.sagepub.com/doi/10.1177/10915818241255885
4. European Commission – *Roadmap towards phasing out animal testing* https://single-market-economy.ec.europa.eu/sectors/chemicals/reach/roadmap-towards-phasing-out-animal-testing_en
5. EMA – *New Approach Methodologies Horizon Scanning Report* https://www.ema.europa.eu/en/documents/report/new-approach-methodologies-eu-horizon-scanning-report_en.pdf
6. UK Government – *Replacing Animals in Science Strategy* https://www.gov.uk/government/publications/replacing-animals-in-science-strategy
7. FDA – *Roadmap to Reducing Animal Testing in Preclinical Safety Studies* https://www.fda.gov/files/newsroom/published/roadmap_to_reducing_animal_testing_in_preclinical_safety_studies.pdf
8. Preclinical Safety Evaluation of Biotechnology-Derived Pharmaceuticals: https://database.ich.org/sites/default/files/S6_R1_Guideline_0.pdf
9. BigPicture: https://bigpicture.eu/
10. eTRANSAFE: https://etransafe.eu/
11. eTOX: https://www.ihi.europa.eu/projects-results/project-factsheets/etox
12. imSAVAR: https://imsavar.eu/
13. VICT3R: https://www.vict3r.eu/
14. ASPIS: https://aspis-cluster.eu/
15. NC3Rs Virtual Second Species - *Applying advanced computational and mathematical modelling approaches to develop a suite of virtual dog tissues and organs to model toxicological endpoints for new chemical entities* https://nc3rs.org.uk/crackit/virtual-second-species
16. Review of the use of two species in regulatory toxicology studies | NC3Rs: https://nc3rs.org.uk/our-portfolio/review-use-two-species-regulatory-toxicology-studies-ichm3#two-species-phase-ii-ich-m3r2
17. Passini et al (2025). OS02-10 Analysis of the use of two species in regulatory toxicology studies for molecules following ICH M3(R2) Toxicology Letters (EUROTOX 2025) https://www.sciencedirect.com/science/article/pii/S0378427425017527
18. Assessment List for Trustworthy Artificial Intelligence (ALTAI): https://digital-strategy.ec.europa.eu/en/library/assessment-list-trustworthy-artificial-intelligence-altai-self-assessment
19. European Directive on the protection of animals used for scientific purposes https://eur-lex.europa.eu/eli/dir/2010/63/oj/eng

Glossary

ACRONYM	MEANING
AI	Artificial Intelligence
ALTAI	Assessment List for Trustworthy Artificial Intelligence
EMA	European Medicines Agency
ERA	European Research Area
EC	European Commission
EU	European Union
FDA	Federal Drug Administration
FTM	Foundation Toxicology Model
GDPR	General Data Protection Regulation
ICH	International Conference on Harmonisation
IHI JU	Innovative Health Initiative
IKAA	in-kind contributions to additional activities
IKOP	in-kind contributions to operational activities
LOAEL	Lowest Observed Adverse Effect Level
NOAEL	No Observed Adverse Effect Level
NAM	New Approach Methodology
SOPs	Standard Operating Procedures
SMEs	Small and Medium-sized Enterprises
UK	United Kingdom