Towards a Virtual Second Species in Toxicology for Drug Development

Overall aim

1. The aim of this Challenge is to apply advanced computational and mathematical modelling approaches including Machine Learning, systems biology, systems toxicology and multi-scale modelling to develop a suite of virtual dog tissues and organs to model toxicological endpoints for New Chemical Entities (NCEs). The model will be built and validated on existing data from repeat dose dog studies and will predict key toxicities likely to be observed in the dog upon long-term dosing for any given NCE. The model will provide an evidence base to support the use of only a single, rodent, species for chronic toxicity studies, with the longer term goal of removing the requirement for an *in vivo* dog chronic toxicity study.

Duration

2. This is a three Phase mega-Challenge with funding for up to five years. Phase 1: nine months, Phase 2: Up to three years, Phase 3 (subject to successful completion of Phase 2): Up to two years.

Budget

3. Phase 1: Up to £100k, Phase 2: Up to £1.6 million, Phase 3: Up to £1 million.

Sponsors

4. Challenge Sponsors define the Challenges in collaboration with the NC3Rs to set out the business case and 3Rs benefits, with a view to using the product developed. Sponsors are required to provide in-kind contributions and/or funding to help solve the Challenges.

5. The Challenge is sponsored by the following companies: **Bayer AG, Eli Lilly and Company, Genentech Inc., Gilead Sciences Inc., GSK, Merck Healthcare KGaA and Roche.**

Partners

6. Challenge partners collaborate with the NC3Rs to provide additional resources to successful applicants to help deliver the Challenge.

7. The Challenge Partners are: **eTransafe - Innovative Medicines Initiative (IMI), Simomics Ltd.**

Co-funders

8. EPSRC

Background

Toxicology in the pharmaceutical industry

9. Animals are used in non-clinical studies to assess the efficacy and potential toxic effects of drugs before their first use in humans and alongside subsequent clinical studies. The safety studies aim to identify target organs of toxicity, assess reversibility of potential effects, and assess exposure-response relationships. The data generated informs the safe human starting dose and clinical monitoring. While first-in-human studies are usually regarded as very safe, not all relevant human toxicities are identified in animal studies (1).

10. Current regulatory guidelines for NCEs such as small molecule drugs usually require safety and tolerability data from two species, a rodent (i.e. rat or mouse) and a non-rodent (i.e. dog, minipig or non-human primate), before administration of potential new medicines in first-in-human clinical trials. Information typically available before these studies are started include early pharmacokinetic (PK) data, physiochemical information, *in vitro* data (e.g. secondary pharmacology off-target profiling and toxicological screening data), early safety pharmacology studies and any specific investigations to evaluate target-related risk or liabilities with high incidence (e.g. liver). Two different species are used for non-clinical safety assessment to account for species-specific differences in susceptibility. Unless an effect is known to be species specific, toxicities seen in one species are regarded as potential human toxicities to be closely monitored in clinical studies, if regarded as manageable. Longer term toxicity studies (including chronic treatment duration of up to 39 weeks) are also conducted in the same two species to support longer clinical studies as well as potential marketing authorisation of drugs for long-term use. There is good evidence that detection of toxicity in at least one of the toxicology study species is indicative of the potential detection of adverse events in clinical observations. However, the absence of toxicological findings in a non-clinical study is not always indicative that no adverse events will be detected in humans and can lead to late stage drug attrition (2,3). The variability in concordance between non-clinical species and adverse events detected in humans may be due to differences in the translation of effects for specific target organs of interest (e.g. good concordance for cardiovascular effects versus poor translation for central nervous system effects such as headache and nausea), the relatively small number of animals and humans studied during early trials and the lack of appropriate non-clinical models.

11. The research and development landscape has changed considerably since these regulatory requirements have been in place, with new *in vitro* and *in silico* technologies available to evaluate safety in addition to the standard *in vivo* approach. Opportunities to refine and reduce animal use within toxicology studies have been explored (4) and some flexibility in guideline testing requirements have been incorporated, for example, chronic toxicity studies are not required for drugs indicated for acute/short-term use, or in life-threatening indications. With emerging technologies, there is now the opportunity to consider the relevance of the animal models used in the development of NCEs.

Mathematical modelling and computational approaches to support a move to the use of a single rodent species in toxicology

12. Work by a NC3Rs-led consortium involving 30 pharmaceutical companies and regulatory bodies from the UK, Europe and North America identified opportunities to use only a single species for chronic toxicity studies (of 13 (sub-chronic) weeks to 26 or 39 (chronic) weeks duration) for a wider range of drug modalities than currently permitted by the regulations (5). This exercise evaluated data retrospectively and additional evidence is required to prospectively determine when a single species chronic toxicity approach may, or may not, be applicable. For NCEs such as small molecule drugs, these opportunities may not be widely adopted until convincing evidence on prediction of outcome of chronic studies is achieved followed by a change in international regulatory guidelines.

13. The pharmaceutical industry is exploiting advances in human-relevant *in vitro* sciences and mathematical modelling, data interrogation and analysis to improve approaches used in the assessment of the safety of new medicines. Organisations in the USA such as the National Toxicology Program and the Food and Drug Administration (6,7) have outlined commitments to support the development of these New Approach Methodologies (NAMs) that have the potential to revolutionise drug development. Advances in mathematical and computational modelling, accelerated by Machine Learning, are poised to revolutionise many aspects of daily life, (8,9) and there is a strong research base exploiting these capabilities to advance drug discovery and development. Integrating data from non-clinical studies with *in silico* tools to maximise the knowledge derived, has the potential to improve the safety assessment of candidate molecules and expedite the drug development process, delivering significant cost benefits and reducing animal use and drug attrition. For example, quantitative structure–activity relationship (QSARs) are widely used in the prediction of physicochemical properties during the selection and optimisation of lead candidates during drug discovery. QSARs have also been applied extensively in toxicology and when combined with read-across and advanced Machine Learning techniques, are increasingly powerful tools that have the potential to provide critical building blocks to parameterise complex biological models (10). However, QSARs are limited by their requirement for adequate knowledge of the chemical space and the extent of their validation (applicability domains), and while useful in screening of novel compounds based on their chemical similarity to historical data, they are not yet able to fully meet regulatory non-clinical testing requirements.

14. Other modelling approaches include integrated PK and pharmacodynamic (PD) modelling that aim to better characterise drug exposure and response relationships. These models can be further expanded into multi-scale models by building in layers of biological complexity from the sub-cellular to the organ system level and even across populations. Classical modelling approaches can be limited by inherently variable and often incomplete biological datasets that need to be accounted for when creating prediction models. Machine Learning can improve the ability of models to handle large, complex datasets, and when combined with the wealth of non-clinical data available and existing physiological models, has the potential to drive a step-change in predictive toxicology (11). The ability of some Machine Learning algorithms to learn and improve over time and their potential to address data gaps where there is adequate coverage from overlapping areas, further improves

their potential predictivity and accuracy. Advances specifically looking to improve models in predictive toxicology include work using neural network approaches to model human organ toxicity and multi-species acute toxicity end points using novel prediction models (12,13). Additionally, agent-based morphometric models of dynamic signalling networks, parameterised by *in vitro* data on human molecular and cellular targets, can represent emergent behaviours and serve as *in silico* chemical screening platforms (14,15). Multi-scale mathematical models combining Machine Learning, agent-based, PK/PD and cell or tissue system models within a quantitative systems toxicology framework also offer the opportunity to represent the interplay between organs in a virtual environment (15,16). With the global bio-simulation market projected to grow to over $9 billion by 2028 (17), there are strong drivers to accelerate the development and accuracy of these models in assessing the safety of new medicines.

15. There is now the potential to apply advanced *in silico* tools and approaches to support the building of a more robust evidence base to facilitate moving towards using a single (rodent) species, without increasing risk to humans. Utilising the vast amount of historical and contemporary dog study data, a virtual model could be used for the assessment of potential target organ toxicities in the dog. Other non-rodents, such as non-human primates (NHPs) and minipigs, are also common non-rodent species used for small molecule testing. However, the historical use of the dog for small molecule development has produced considerable data in the literature and study reports that can be exploited. There may be potential to expand the model to develop other virtual non-rodent species in the future, and to other drug modalities.

16. A virtual human in which to assess potential toxicity would perhaps be the ideal model to develop, however, this is likely to be too big of a translational step (from rodent directly into human, without some level of non-rodent data) for acceptance by clinicians and regulators. The validation required for a model predicting human effects will also be much more difficult to achieve than acceptance of a model used to predict long-term dog study outcomes (where there may be the possibility to also validate predictions with shorter term dosing studies in dog and/or rat). Dogs used in toxicology studies are genetically less heterogeneous than the human population providing data more amenable for modelling purposes.

The Challenge

17. This Challenge aims to develop a model or suite of models that will ultimately replace the use of the dog in six to nine-month (chronic) toxicology studies. Proposals that aim to tackle the problem using a system-wide approach are welcomed. It is, however, acknowledged that the early and mid-term aims may focus on the development of a model that can identify the potential toxicity risk for a series of key target organs known to be affected in the dog, or that have been documented to be likely indicators of risk in humans.

18. The model is targeted towards the identification of toxicities that may develop upon long-term dosing in the dog, but it could also be used early within discovery as a screen for potential ‘show-stopper’ toxicities in the dog and focus investigations towards those target organs of concern (using other *in*

vitro and *in silico* models) before testing in animals. Replacement of the dog in chronic toxicity testing will require significant validation for regulatory acceptance and there will likely be a period where results from this model would be run in parallel with the current regulatory testing requirements. The model developed through this Challenge will contribute to the growing evidence base and capability of NAMs to provide robust and predictive methods to assess toxicities and reduce the reliance on animals.

19. Rodent data will be available to support model development where required. Dog data is generally evaluated in conjunction with rodent data for small molecule safety programmes, and predictivity of the model in relation to associated rodent data is an important factor when developing the model.

3Rs benefits

20. In the UK in 2020, of the experimental procedures for repeat dose toxicity[1] , there were 10,670 using mice, 27,432 using rats, 2,082 using dogs and 1,142 using non-human primates.

21. The typical design for a dog chronic toxicology study is to dose animals daily for six or nine months, dependent on the region where data will be submitted and reviewed (Europe or USA/Japan respectively (18)). It is common to include four main test groups (control and three dose levels) of three or four male and female animals per group, plus additional groups to assess recovery from any effects (usually restricted to control and high dose, using two male and female animals per group) resulting in 40 dogs typically being used for each chronic toxicity study, with each animal undergoing clinical assessment and repeated blood sampling over specific time periods.

Industry benefits

22. The cost of a typical chronic toxicity study is often in excess of £500k, not including the additional costs associated with compound production and quality certification, formulation methodology support and data review. If this study could be replaced with a modelling tool, there would be a significant cost and time saving for industry, as well as the potential for the generation of more informative data.

23. With the time taken to develop new medicines ranging from 10 to 15 years, advances in NAMs such as *in silico* models have the potential to deliver improved decision-making tools, with better mechanistic understanding that result in more rapid discovery and development of medicines.

[1] In the UK, the number of animals used in scientific procedures is provided annually by the Home Office. Although the UK figures specify the number and species of animals used, their use for repeated dose toxicity covers a broader range of toxicity tests than described for this Challenge.

Challenge partners

eTRANSAFE

24. Over the past decade, collaborative industry projects have collected a significant amount of non-clinical data from toxicology and safety assessment studies, providing resources of curated databases of study data from partner pharmaceutical companies. These databases have then been used to generate novel tools for predictive toxicological modelling (19, 20, 21). One current project – Enhancing TRANslational SAFEty Assessment through Integrative Knowledge Management (eTRANSAFE), building on the previous project –eTox[2], is developing a technological architecture for data sharing of non-clinical and clinical data, with in depth data integration and data exploitation capabilities. This effort is creating a database of curated SEND ("standards for the exchange of non-clinical data") reports provided by 12 partner pharmaceutical companies in addition to what was generated during the preceding eTox initiative and combining it with chemoinformatics, bioinformatics and clinical drug safety data to assess the predictive translation of non-clinical data to humans. The project has already developed the potential for a 'rat virtual control group' (22) which could reduce the need for concurrent controls in each study, instead making comparisons of test data with the historical database. The wealth of data that has been collected and curated through eTox and eTransafe provides a significant resource for delivery of this Challenge, and the activities developed through this Challenge will add additional outcomes and impacts to the core eTRANSAFE goals. eTRANSAFE are partnering with the NC3Rs to support data provision for the successful Challenge winners.

Simomics

25. Software technologies from Simomics help support transparency for *in silico* models, embedded in their virtual laboratory approach (23). Components of their core technology were developed through previous CRACK IT funding and Simomics are keen to offer this to the successful teams, if required, to help increase transparency of the models developed, aiding uptake and acceptance.

26. Simomics will offer free, non-commercial, licences (for the purposes of this project) for a number of their software tools, along with reasonable technical support, depending on the needs of awardees. Tools include: the ability to annotate of software models with assumptions and the rationale for design, the ability to trace the reliance of models on key evidence and drawing tools that allow the creation of arguments to support the design and rationale for models and results. These tools will be offered to successful applicants as an option to help build their models but are not required.

[2] Integrating bioinformatics and chemoinformatics approaches for the development of expert systems allowing the *in silico* prediction of toxicities

Key deliverables

27. The aim of this Challenge is to develop advanced mathematical and computational in silico models that will contribute to reducing or eliminating the need for chronic toxicity studies in the dog during drug development. The model(s) should be able to both predict chronic toxicological effects of small molecules in key target organs of the dog not detected in the four-week study and any changes in severity of those that were detected. Its use in the first instance would be supported by sub-chronic dog (and rodent if required) toxicology study data in order to build confidence before further development with the longer term aim (not within scope of the Challenge) to remove the requirement for use the dog as a second species for NCEs. The model should permit understanding of the mechanism of action and where possible provide a prediction of the quantitative or qualitative extent of the effect.

28. Delivery of this Challenge should:

- Permit the toxicity liability assessment of low molecular weight NCEs such as small molecules, peptides and oligonucleotides that use two species in development. Biologics and other drug modalities that may use two species are not in scope.
- Include models that are either multi-organ and/or multi-system and deliver a step-change that advances current scientific and technological approaches to predict toxicity.
- Include provision of evidence of appropriate data management and cybersecurity processes to ensure any confidential data provided is stored and managed securely.
- Include strong project management processes to ensure timely delivery of milestones.

29. Delivery of this Challenge should not:

- Focus on solely the creation or further development of QSAR models.
- Involve the generation of new *in vitro* models.

Phase 1 deliverables

30. The deliverables for Phase 1 are:

- Develop a proof-of-concept approach that demonstrates how you would address the wider Challenge.
- Using control dog (and rodent where required) data provided, develop virtual *in silico* biology framework for the model.
- Provide evidence of the effectiveness of the approach (e.g. initially by focusing on a single target organ/system and plans for how this would be extended or combined with other target organs/systems).
- Identify additional databases or other open sources of data for use within the project.

- Propose approaches to curate and extract data from other sources in Phase 2 (see Phase 2 in-kind contributions, e.g. legacy pdf reports).
- Identify additional information needed that the Sponsors may be able to provide in Phase 2.

Phase 2 deliverables

31. Phase 2 includes essential and desirable deliverables. **Essential deliverables** are:

- Full development of a system-wide or multiple connected organ-system models that will permit the interrogation and prediction of toxicological events.
- Determination of the prediction accuracy (sensitivity and specificity) of the model to dog toxicities as seen in four-week studies, both positives and negative toxicological outcomes and for a suite of target organs.
- Demonstration of the prediction accuracy (sensitivity and specificity) of the model against a series of compounds with known animal and/or human toxicological endpoints for chronic studies.
- The model should identify toxicities including organ system affected, toxicological endpoint and where possible, mechanism of action.

32. As part of the **essential deliverables**, contractors must:

- Develop the prototype end-user interface in close collaboration with the Sponsors to ensure it is fit-for-purpose.
- Incorporate and demonstrate use of best/industry standard computational modelling practices.
- Ensure that use of pre-existing software is free of third-party restrictions for use.
- Develop plans for how the model will be made widely available upon Challenge completion.

33. The Challenge also includes desirable deliverables that are intended to maximise the use of the data provided and the ambition for the model in the longer term. From a baseline of four-week (sub-chronic) toxicity data in two species, toxicities may either diminish or resolve, progress in severity, or new target organ toxicities may be identified upon chronic dosing. Where possible, the model should:

- Predict any additional toxicities caused by longer term dosing and not identified in four-week dog studies.
- Model changes in severity of any toxicities identified in four-week studies.
- Predict the effects, including potential toxicities, of drug metabolites.
- Demonstrate the potential to predict and model multi-organ toxicity interactions such as blood count linked with bone marrow toxicity, or steroid imbalance causing reproductive organ toxicity.

- Demonstrate the potential for the ability to monitor dose response relationships over time.

Phase 3 deliverables

34. Phase 3 awards are subject to assessment on the delivery of Phase 2. The Phase 3 deliverables are:

- Development of an accurate system to establish the relationship between predicted toxicity and drug exposure.
- Identification of potential for cross-species translation between rodent, dog and human.
- Production of clear end-user documentation.
- Iterative testing of the model with new compounds and datasets.
- Demonstration of iterative testing with end-users.
- Incorporation of end-user requirements to control experimental parameters.
- Consideration and discussion of the new approach with regulatory agencies and pharmaceutical networks.
- Initial deployment of the model to early end-users and development of strategies for wider dissemination.

In-kind contributions

35. The Challenge will be supported by access to historical dog toxicity data supplied by sponsoring and collaborating companies that where possible, will be well-curated:

Phase 1 in-kind contributions

- Control study data and a limited number of sub-chronic and chronic control and treated data for the dog provided by the eTransafe IMI consortium.
- Discussion and toxicological expert input in proposed approach

Phase 2 in-kind contributions

- Sub-chronic and chronic control and treated data study data provided by the eTransafe IMI consortium.
- Sub-chronic and chronic control and treated study data and other (e.g. DMPK reports) for the dog provided by the Sponsors.
- Sub-chronic and chronic control and treated data and other (e.g. DMPK reports) for the rodent provided by the Sponsors where required and available.
- Discussion and toxicological expert input in model use requirements.

Phase 3 in-kind contributions

- Phase 3 awards are subject to assessment on the delivery of Phase 2.
- Additional compound data for testing from Sponsors and additional partners.
- Promotion of model through partner networks to engage partners to support further testing and deployment.

Additional publicly available toxicology data

36. It is expected that alongside the data provided by the Sponsors and Partners, applicants will utilise publicly available data sets, adverse outcome pathways and research. A growing number of resources and databases are available, and examples can be found here:

- Gopal P., et al. (2019) In Silico Toxicology Data Resources to Support Read-Across and (Q)SAR. Frontiers in Pharmacology, 10:561. doi:0.3389/fphar.2019.00561.
- Jain S., et al. (2021) Large-Scale Modeling of Multispecies Acute Toxicity End Points Using Consensus of Multitask Deep Learning Methods. Journal of Chemical Information and Modeling. 61(2):653-663. doi: 10.1021/acs.jcim.0c01164.
- The Comparative Toxicogenomics Database http://ctdbase.org/ .

References

1. Ponzano S., *et al.* (2018). Promoting Safe Early Clinical Research of Novel Drug Candidates: A European Union Regulatory Perspective. Clinical Pharmacology Therapeutics 103: 564-566. https://doi.org/10.1002/cpt.899.

2. Monticello TM., *et al.* (2017) Current non-clinical testing paradigm enables safe entry to First-In-Human clinical trials: The IQ consortium non-clinical to clinical translational database. *Toxicology and Applied Pharmacology*. 1;334:100-109. doi: 10.1016/j.taap.2017.09.006

3. Sanz, F., *et al.* (2017). Legacy data sharing to improve drug safety assessment: the eTOX project. *Nature Reviews Drug Discovery* 16(12): 811-812. doi.org/10.1038/nrd.2017.177

4. Sewell F., *et al.* (2016). Opportunities to apply the 3Rs in Safety Assessment programs. *ILAR journal* 57: 234-245. doi.org/10.1093/ilar/ilw024

5. Prior H., *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* 113: e104624. doi: 10.1016/j.yrtph.2020.104624

6. A Strategic Roadmap for Establishing New Approaches to Evaluate the Safety of Chemicals and Medical Products in the United States

7. Avila A.M., *et al*. 2020 An FDA/CDER perspective on non-clinical testing strategies: Classical toxicology approaches and new approach methodologies (NAMs). *Regulatory Toxicology and Pharmacology*. Volume 114, 104662. doi.org/10.1016/j.yrtph.2020.104662.

8. Transforming our world with AI: UKRI's role in embracing the opportunity

9. AI Council AI Roadmap

10. Luechtefeld T., *et al.* (2018). Machine Learning of Toxicological Big Data Enables Read-Across Structure Activity Relationships (RASAR) Outperforming Animal Test Reproducibility. *Toxicological Sciences* 165(1), 198-212. doi.org/10.1093/toxsci/kfy152

11. Wenzel J., Matter, H., and Schmidt, F. (2019). Predictive Multitask Deep Neural Network Models for ADME-Tox Properties: Learning from Large Data Sets. *Journal of Chemical Information and Modeling*, 59 (3), 1253-1268 doi: 10.1021/acs.jcim.8b00785

12. Sosnin S., *et al.* (2019). Comparative Study of Multitask Toxicity Modeling on a Broad Chemical Space. *Journal of Chemical Information and Modeling 59* (3), 1062-1072 doi: 10.1021/acs.jcim.8b00685

13. Jain S., *et al.* (2021) Large-Scale Modeling of Multispecies Acute Toxicity End Points Using Consensus of Multitask Deep Learning Methods. *Journal of Chemical Information and Modeling*. 61(2):653-663. doi: 10.1021/acs.jcim.0c01164

14. Saili KS., *et al*. (2019) Systems Modeling of Developmental Vascular Toxicity. *Current Opinion in Toxicology*. 15(1):55-63. doi:10.1016/j.cotox.2019.04.004

15. Kleinstreuer N., *et al*. (2013) A Computational Model Predicting Disruption of Blood Vessel Development. *PLOS Computational Biology* 9(4): e1002996. doi.org/10.1371/journal.pcbi.1002996

16. Ferreira S., *et al*. (200) Quantitative Systems Toxicology Modeling To Address Key Safety Questions in Drug Development: A Focus of the TransQST Consortium. *Chemical Research in Toxicology 33* (1), 7-9. doi: 10.1021/acs.chemrestox.9b00499

17. Global Biosimulation - Market and Technology Forecast to 2028

18. ICHM3(R2) June 2009. Non-clinical safety studies for the conduct of human clinical trials and marketing authorization for pharmaceuticals.

19. Steger-Hartmann, T. *et al*. (2017). "The IMI eTOX initiative Data mining, read-across and predictive models for target evaluation and early drug candidate assessment." *Toxicology Letters* 280: S19-S19. https://doi.org/10.1016/j.toxlet.2017.07.036

20. Clark M, Steger-Hartmann T. (2018). A big data approach to the concordance of the toxicity of pharmaceuticals in animals and humans. *Regulatory Toxicology and Pharmacology* 96:94-105. doi: 10.1016/j.yrtph.2018.04.018.

21. Pognan F., *et al.* (2021). The eTRANSAFE Project on Translational Safety Assessment through Integrative Knowledge Management: Achievements and Perspectives. *Pharmaceuticals* 14, 237: 1-18. doi.org/10.3390/ph14030237

22. Steger-Hartmann T., *et al.* (2020). Introducing the concept of virtual control groups into preclinical toxicology testing. *ALTEX* 37(3):343-349 doi/10.14573/altex.2001311

23. Simomics