# ICH S5(R3)

## Detection of developmental and reproductive toxicity for human pharmaceuticals

*Step 4* document – to be implemented

18 February 2020

International Council for Harmonisation of Technical Requirements for Pharmaceuticals for Human Use



## Legal Notice

## Background

- This document has been signed off as *Step 4* document (Feb 18th 2020) to be implemented by the ICH Regulatory Members
- This document was developed based on a Concept Paper (March 27th 2015)





## Key Principles 1/2

- The aim of developmental and reproductive toxicity (DART) studies is to reveal any effect of the pharmaceutical on mammalian reproduction and development relevant for human risk assessment.
- The risks to all stages of reproduction and development should be assessed, unless the stage is not relevant to the intended population.

## Key Principles 2/2

- **The 3rd revision**
  - brings the guideline into alignment with other ICH guidelines.
  - elaborates on the use of exposure margins in dose level selection.
  - incorporates a section on risk assessment.
  - expands the scope to include vaccines and biopharmaceuticals.
  - describes qualification and use of alternative assays (AAs).
- **A maintenance procedure is applied to Annexes 1 (in vivo study designs) and 2 (Alternative Assays) aiming to facilitate proper implementation of novel testing paradigms and regulatory acceptance of alternative assays supporting global 3R (replacement, reduction, refinement) efforts.**





## Guideline Objectives

- **Recommendation of international standards for the assessment of nonclinical DART testing supporting clinical trials and marketing authorization.**
- **Description of potential strategies and study designs to supplement available data to identify, assess, and convey risk.**

## Scope of the Guideline

**Guideline applies to:**

Pharmaceuticals, including biotechnology-derived products

Vaccines (and their novel constitutive ingredients) for infectious diseases

Novel excipients

**Does not apply to:**

cellular therapies, gene therapies and tissue-engineered products.

**Whether and when non-clinical studies are warranted is determined by ICH M3(R2), ICH S6(R1), and ICH S9.**





## Table of Contents


- **General Considerations on Reproductive Toxicity Assessment**
- **Design and Evaluation of *in vivo* Mammalian Studies**
- **Test System Selection**
- **Dose Level Selection, Route of Administration and Schedule**
- **Possible Combination Study Designs in Rodents**
- **Data Reporting and Statistics**
- **Principles of Risk Assessment**

## General Considerations on Reproductive Toxicity Assessment 1/7

**The following stages of reproduction should be assessed, if relevant to the intended population:**

A) Premating to conception

B) Conception to implantation

C) Implantation to closure of the hard palate

D) Closure of the hard palate to the end of pregnancy

E) Birth to weaning

F) Weaning to sexual maturity

**Timing of studies is dependent on study populations and phase of pharmaceutical development (see ICH M3, ICH S6 and ICH S9).**





## General Considerations on Reproductive Toxicity Assessment 2/7

**Stages covered in individual studies are left to the discretion of the Sponsor**

**Typically three *in vivo* study types have been evaluated:**

1) a fertility and early embryonic development study (FEED - stages A and B),

2) embryo-fetal development studies in two species (EFD - stages C and D), and

3) a pre- and a postnatal development study (PPND – stages C through F).

## General Considerations on Reproductive Toxicity Assessment 3/7

**Key factors to consider when developing an overall integrated testing strategy:**

- The targeted patient population and conditions of use;
- The formulation of the pharmaceutical and route(s) of administration intended for humans;
- Relevant data on toxicity, pharmacodynamics, pharmacokinetics, and pharmacological similarity to other pharmaceuticals;
- Aspects of the general biology of the pharmaceutical target, or known roles of the target in reproduction or development.





## General Considerations on Reproductive Toxicity Assessment 4/7

**Target Patient Population/ Therapeutic Indication**

…can influence the extent of DART testing.

Studies evaluating all stages of reproduction and development are not warranted if the disease indicates that DART will have minimal impact on the risk of the pharmaceutical in the target population.

For example, studies covering all stages are not necessarily appropriate for an exclusively post-menopausal female patient population, for use in the pediatric or juvenile pre-pubescent population, or for patient populations in hospitalized settings where pregnancy can be excluded.

## General Considerations on Reproductive Toxicity Assessment 5/7

### Pharmacology Considerations

Check compatibility of the intended pharmacology with fertility, normal EFD, or assessment of particular endpoints (e.g., a general anesthetic and assessment of mating behavior).

- Known effects of compounds with similar pharmacology.
- Known effects of target engagement.
- Known effects on humans with related genetic diseases.





## General Considerations on Reproductive Toxicity Assessment 6/7

### Toxicity Considerations

Always consider existing toxicology data for the compound concerning dose levels, toxicokinetics, dosing duration.

Repeated-dose toxicity studies (RDTS) with sexually mature animals can provide important information.

### Timing Considerations

For safe use of the pharmaceutical in clinical trials or the intended patient population consult ICH M3, ICH S6, and ICH S9.

Under limited cirumstances results of properly qualified Alternative Assays (*in vitro*, *ex vivo*, and non-mammalian *in vivo* assays ) have the potential to defer or replace conventional *in vivo* studies.

## General Considerations on Reproductive Toxicity Assessment 7/7

**Toxicokinetics (TK)**

Exposure data (see ICH S3A) can be generated in either reproductive (DRF or pivotal) or RDTSs

- It is recommended to determine if pregnancy alters exposure.
- GLP-compliant TK data in pregnant animals is expected for exposure-based dose selection.

The concentration in the embryo/fetus can facilitate data interpretation.

Evidence of lactational excretion can be shown by sampling milk or demonstration of exposure in offspring during the pre-weaning period.





## Design and Evaluation of *in vivo* Mammalian Studies 1/10

**Should allow for evaluation of all stages of the reproductive process**

- In some species (e.g., NHP) it is not possible to evaluate all stages
- Typically the 3-study design is applied

Combinations of these study designs can be conducted to address specific product needs and to reduce animal use.

Annex 1 contains study details for the FEED, EFD, and PPND studies, and combinations thereof.

ICH S5(R3) – Step 4

## Design and Evaluation of *in vivo* Mammalian Studies 2/10

**Fertility and Early Embryonic Development (FEED)**

To examine stages A and B, typically only in rodents.

Results from RDTSs of ≥ 2 weeks can replace dose range finding (DRF) for fertility studies.

Determine the affected sex for AEs on fertility (separate treatment arms).

Reversibility of adverse effects can have an important impact on risk assessment.



ICH S5(R3) – Step 4

## Design and Evaluation of *in vivo* Mammalian Studies 3/10

**Fertility and Early Embryonic Development (FEED) – contd.**

**FEED study female rodents:** effects on the estrous cycle, tubal transport, implantation, and development of preimplantation stages
**FEED study design male rodents:** effects on spermatogenesis and epididymal transport typically studied after 2 to 4 wks of treatment prior to cohabitation

Assessment of the full spermatogenic cycle and epididymal transport can be appropriate for toxicities to the testis suggested from RDTS.

FEED study permits detection of functional effects otherwise not detectible by histological examinations of the male reproductive organs.

Consider additional examinations integrated into RDTS and/or fertility studies.

## Design and Evaluation of *in vivo* Mammalian Studies 4/10

**FEED - Considerations for Biopharmaceuticals**

Use rodents or rabbits if it is pharmacologically active in one of these species.

Mating evaluations are not generally feasible in non-rodents such as dogs and NHPs.

Histopathological examinations of the reproductive tissues from the RDTSs studies of at least three months duration can serve as a substitute for the fertility assessments.

- Animals should be sexually mature at study initiation.
- No functional assessment of fertility can be made.





## Design and Evaluation of *in vivo* Mammalian Studies 5/10

**Embryo-Fetal Development (EFD)**

include evaluation of Stages C through D following Stage C exposure

Small molecules: evaluated in rodent and non-rodent species

- At least one of the test species should exhibit the desired PD. Otherwise consider non-routine species, genetically modified animals, or surrogates.
- Genetically modified animals/surrogate most useful for hazard identification.
- In absence of relevant models, two species testing should be conducted to detect the adversity of off-target effects or secondary pharmacology.

ICH S5(R3) – Step 4

## Design and Evaluation of *in vivo* Mammalian Studies 6/10

**Embryo-Fetal Development (EFD) – contd.**

Under some circumstances, it may not be necessary to conduct EFD studies in multiple species. These include:

- Clearly positive results in a single species at exposures similar to that at the projected clinical exposure at the maximum recommended human dose (MRHD)
- Evidence suggesting an adverse effect of the intended pharmacological mechanism on EFD (e.g., MoA)

Under limited circumstances, alternative assays can be used



ICH S5(R3) – Step 4

## Design and Evaluation of *in vivo* Mammalian Studies 7/10

**EFD – Considerations for Biopharmaceuticals**

Typically assessed in one rodent and one non-rodent species if pharmacologically relevant

- If the rodent is not pharmacologically relevant, conduct single species EFD testing in pharmacologically relevant non-rodents.
- If NHP is the only relevant species, consider an ePPND study.
- Single relevant species is sufficient for treatment of advanced cancer (ICH S9).
- No relevant species: consider surrogates/transgenic models (ICH S6).
- If there are no relevant species, genetically modified animals or surrogates available, *in vivo* reproductive toxicity testing is not meaningful. The approach used for risk assessment should be justified.

## Design and Evaluation of *in vivo* Mammalian Studies 8/10

**Alternative Approaches for Addressing EFD Risk**

Are encouraged and have the potential to defer or replace conventional *in vivo* studies and reduce animal use

Should provide a level of confidence for human safety assurance at least equivalent to that provided by the current testing paradigms





## Design and Evaluation of *in vivo* Mammalian Studies 9/10

**Approaches to Defer Definitive *In Vivo* Testing as Part of an Integrated Testing Strategy** (not needed for US)

additional options to ICH M3(R2) compliant pEFD study results from two species

1) Qualified alternative assays which predict the outcome in one species can be combined with a pEFD from a second species to enable the limited inclusion of WOCBP (up to 150 WOCBP for up to 3 months). The alternative assay and the second species should generally cover both a rodent and a non-rodent species.

2) Additional endpoints incorporated into at least one GLP pEFD study (with increased group size and including fetal skeletal examinations) in a pharmacologically relevant species, if available, combined with a pEFD study in a 2nd species to include an unlimited number of WOCBP in clinical trials through Phase 2.

## Design and Evaluation of *in vivo* Mammalian Studies 10/10

**Pre- and Postnatal Development (PPND)**

- Include evaluation of Stages C through F (delayed effects)
- Usually in rodents, other species are possible
- Usually prior studies inform about study design and dose levels, otherwise consider preliminary PPND studies
- Modified PPND/ePPND study designs ☞ see ICH S11

**PPND - Considerations for Biopharmaceuticals**

In ePPND studies with NHP it is not generally feasible to follow the offspring through maturity





## Test System Selection 1/5

**Routine Test Species**

Mammalian, well-characterized species, relevant for the endpoints

Advantages to use the same species and strain as in already completed toxicity studies (PK, metabolism, DRF).

Rats are the most often used primary species for all stages of reproductive testing (practicality, general knowledge of pharmacology, background data)

For EFD only, rabbits are typically used as a second mammalian non-rodent species

## Test System Selection 2/5

***Species Selection for Preventative and Therapeutic Vaccines against Infectious Diseases***

- Should demonstrate an immune response
- Typically single species (rabbits, rats, or mice) testing applying a full human, single dose level can be sufficient
- When there is a lack of an appropriate animal model (including NHP), an EFD toxicity study in rabbits, rats, or mice can still provide important information





## Test System Selection 3/5

**Non-routine Test Species**

Should be considered when use of routine species is not appropriate.

Refer to Annex 1 for advantages and disadvantages of different species.

Includes NHPs, typically – if pharmacologically relevant - used for EFD studies and early postnatal development assessment of biopharmaceuticals (ICH S6).

## Test System Selection 4/5

**Use of Disease Models, Genetically Modified Models, and Surrogate Molecules**

Can be valuable for investigating pharmacological effects on development and reproduction

- If data obtained from healthy animals could be misleading
- Should be pharmacologically relevant and appropriate for the endpoints being assessed
- Pathophysiology and animal-to-animal variability should be characterized, while some differences in pathophysiology to humans are acceptable if these are unlikely to confound data interpretation
- If limited historical data, reference data for study endpoints should be available or generated





## Test System Selection 5/5

**Use of Disease Models, Genetically Modified Models, and Surrogate Molecules – contd.**

Genetically modified models can be useful

- to provide information about on-target effects on DART parameters,
- to inform about a functional link between target and adversities observed in routine test species.

In absence of adequate activity against the target in the routine species, surrogate molecules can be used.

## Dose Level Selection, Route of Administration and Schedule 1/7

**Dose Selection**

– *Toxicity–based Endpoint*

Induces a minimal level of toxicity in the parental animals at the high dose. Factors limiting the high dose include:

- Alterations in body weight (gain or absolute; either reductions or increases).
- Exaggerated pharmacological responses.
- Toxicological responses.

– *Saturation of Systemic Exposure Endpoint*

Measured by systemic availability.





## Dose Level Selection, Route of Administration and Schedule 2/7

**Dose Selection – contd.**

– *Exposure Margin Based Endpoint*

25-fold the exposure in pregnant animals at the MRHD are generally considered appropriate as the maximum dose

- Margin established in a GLP-compliant DRF/pEFD or definitive study.
- Usually based on parent drug levels (metabolites see ICH M3 and ICH M3 Q&A).
- Special considerations for prodrugs.
- Consider higher doses if PD activity in the test species is limited at 25-fold.
- GLP-compliant TK data from pregnant animals are expected. Justify the choice for the use of total vs. fraction unbound pharmaceutical exposures (ICH S3A).

## Dose Level Selection, Route of Administration and Schedule 3/7

**Dose Selection - contd.**

– *Exposure-based Approach for Biopharmaceuticals*

Either the maximum intended pharmacological effect in the preclinical species or an approximately 10-fold exposure multiple over the maximum exposure to be achieved in the clinic, whichever is higher.

Consider dose adjustment for differences in target binding affinity and other relevant factors (ICH S6).





## Dose Level Selection, Route of Administration and Schedule 4/7

**Dose Selection – contd.**

– *Maximum Feasible Dose (MFD) Endpoint*

If the physicochemical properties of the pharmaceutical (or formulation) limit the administrable amount of the pharmaceutical.

Maximize exposure, rather than maximize the administered dose (ICH M3 Q&A).

– *Limit Dose Endpoint*

1 g/kg/day, if other dose selection factors have not been attained with lower dose levels (ICH M3).

## Dose Level Selection, Route of Administration and Schedule 5/7

**Dose Selection – contd.**

– *Selection of Lower Dose Levels*

- Desirable to establish a NOAEL for DART, together with dose-response relationship.
- Generally the low dose should provide a low multiple (e.g., 1 to 5-fold) of the human exposure at the MRHD.
- Justify dose levels that yield sub-therapeutic exposures.





## Dose Level Selection, Route of Administration and Schedule 6/7

**Route**

In general the clinical route should be used.

Different routes should be considered if clinical route is not feasible or yields poor exposure.

A single route (with sufficient exposure to the API and relevant metabolites) in the test species can be adequate, if multiple routes are being evaluated in humans.

**Schedule**

Generally mimicking the clinical schedule, but alterations can be justified to ensure adequate exposure.

## Dose Level Selection, Route of Administration and Schedule 7/7

**Dose Selection and Study Designs for Vaccines**

Covers vaccines (adjuvanted or not) used in both preventative and therapeutic indications against infectious diseases.

- The principles apply to nonclinical testing of vaccines for other indications (e.g., cancer).
- Generally not warranted if developed for neonates, pre-pubertal children, or geriatric populations.
- Vaccination regimen should maximize maternal antibody titers.
- Episodic dosing rather than daily dosing (overexposure!) with at least one dose during pregnancy is recommended.
- Consider additional testing strategies for novel, active constituent ingredients (including novel adjuvants).





## Possible Combination Study Designs in Rodents

Majority of programs combine three separate study designs to assess all stages: FEED, EFD, PPND

Various combinations can be conducted to reduce animal use.

A combined Fertility/EFD study plus separate PPND study is common.

Combination of RDTS and fertility study (see Annex 1) in cases

- where no effects on male or female fertility are anticipated, or
- where extending the dosing period is appropriate due to observation of reproductive organ toxicity in a RDTS (no. of mating pairs should ≥16).
- Dosing of females can be extended until the end of organogenesis, thereby allowing evaluation of EFD endpoints.

## Data Reporting

Values should be tabulated in a clear, concise manner.

Apply industry-harmonized terminology to describe fetal morphologic abnormalities.

Reference data for the study endpoints should be available or should be generated during the study to aid data interpretation.





## Statistics

Statistical testing is expected in definitive studies.

Cesarean, fetal and postnatal data summary statistics should be calculated using the litter as the unit of analysis.

Statistical significance need not convey a positive signal, nor lack of statistical significance impute absence of effect.

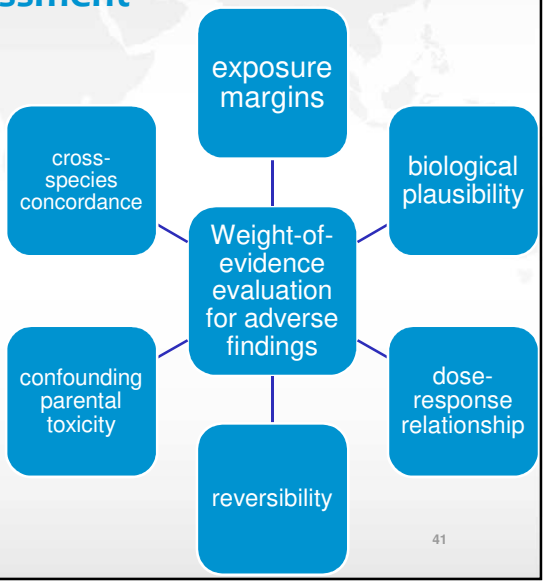
## Principles of Risk Assessment

Should include all available data.

Address any limitations, uncertainties and data gaps.

Definitive *in vivo* studies carry more weight than those from alternative assays or preliminary studies.

For rare malformations, the absence of increased frequency with dose does not always alleviate concern.








## Principles of Risk Assessment

Comparison of exposures at NOAEL and the MRHD is an important component of the risk assessment:

In general, there is increased concern when the NOAEL occurs at exposures less than 10-fold the human exposure at the MRHD;

Effects that are limited to occurrence at more than 25-fold the human exposure at the MRHD are usually of minor concern for the clinical use of the pharmaceutical.

The most relevant margin is generally the exposure metric in the most sensitive species.

## Annex 1: *In vivo* Study Designs

Principle Advantages and Disadvantages of Various Species for DART Testing ☞ Table 1

**Considerations**

Animals should be of comparable age, weight and parity at the start

Typically 16 to 20 litters for rodents and rabbits

- < 16 litters inconsistent results, > 20 to 24 litters per group, consistency and precision is not greatly enhanced.
- If groups are subdivided for different evaluations, the number of animals starting the study should be adjusted accordingly.





## Annex 1: *In vivo* Study Designs

***Fertility and Early Embryonic Development (FEED) Study***

***EFD Toxicity Study***

***PPND Toxicity Study and ePPND Toxicity Study in NHPs***

***Combination Studies***

incorporating appropriate endpoints into a single study

- ***FEED and EFD:*** evaluation of stages A through D, mostly rodents
- ***Male Fertility and Repeated-Dose Toxicology Study***

  To evaluate male fertility during a rodent RDTS, while female fertility and other FEED endpoints need to be evaluated in a separate study.

## Annex 2: Alternative Assays

Under limited circumstances qualified alternative assays can be utilized to support hazard identification and risk assessment.

Potential uses can include:

- Circumstances where there is evidence suggesting an adverse effect on EFD.
- Toxicity in animal species precludes attaining systemic exposures relevant to the human exposures under conditions of use.
- As support for a weight of evidence assessment of equivocal findings.
- As partial support for clinical trials including up to 150 WOCBP for up to 3 months duration.
- Pharmaceuticals being developed for certain severely debilitating or life-threatening diseases or late-life onset diseases.





## Annex 2: Alternative Assays

- Incorporation of these assays into an integrated testing strategy should be justified.
- In accordance with GLP and qualified for context of use (i.e. applicability domain and regulatory conditions under which assay results are reliable).
- Should include drug metabolites (ICH M3).
- No specific assays are recommended, but basic scientific principles are included to assist in assay qualification for regulatory use.
- Alternative assays used to explore mechanism of action are not expected to be qualified in this rigorous manner.

## Annex 2: Alternative Assays

### Qualification of AAs for Prediction of Malformations or Embryo-Fetal Lethality (MEFL)

Should include:

- A thorough description and justification of the predictive model.
- An evaluation of the biological plausibility of the model.
- An assessment of the accuracy and ability for the alternative assay to detect MEFL.
- Definition and justification of the threshold for molecular and metabolic markers predicting MEFL.
- The details of the algorithm employed for determining positive and negative outcomes *in vivo*.





## Annex 2: Alternative Assays

### Qualification of AAs for Prediction of MEFL – contd.

- The list of compounds in each of the training sets and test sets for qualification of the assay and the basis for selection.
- Data sources for all *in vivo* exposure and MEFL data
- The test method's performance for its context of use
- The sensitivity, specificity, positive and negative predictive values, and reproducibility of an assay or battery
- If more than one assay is conducted, a separate description of the performance of each assay, in addition to the integrated assessment used for the predictive model. A clear description of individual data integration.
- Historical data for assay development and use (e.g., viability, numbers and types of malformations), including positive controls.

# Annex 2: Alternative Assays

## Examples of EFD Testing Strategies Utilizing AAs

### *Potential Approach to Defer In Vivo Testing as Part of an Integrated Testing Strategy*

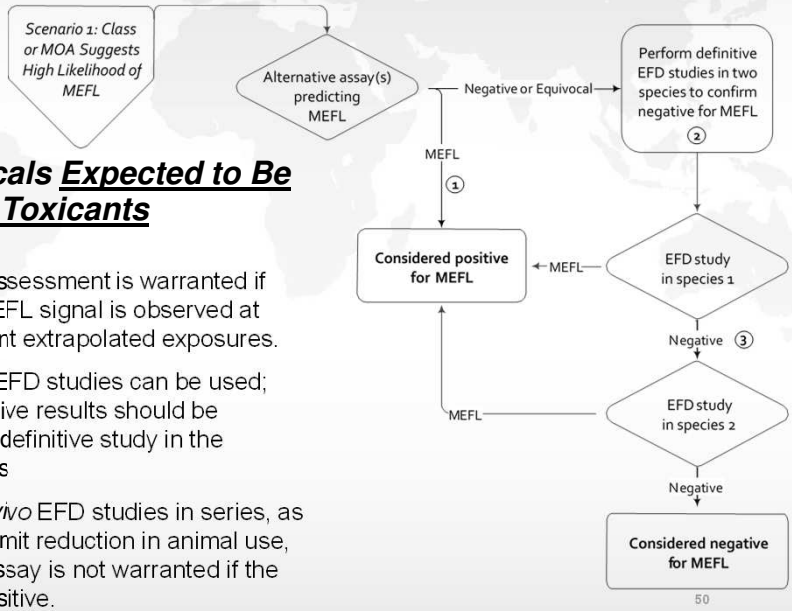
See Section 4.2.3 of the Guidance.

### *Pharmaceuticals Expected to Be Embryo-fetal Toxicants*

Confirmation of expected EF toxicity (mechanism of action, pharmacologic class, target biology) by a qualified alternative assay(s)





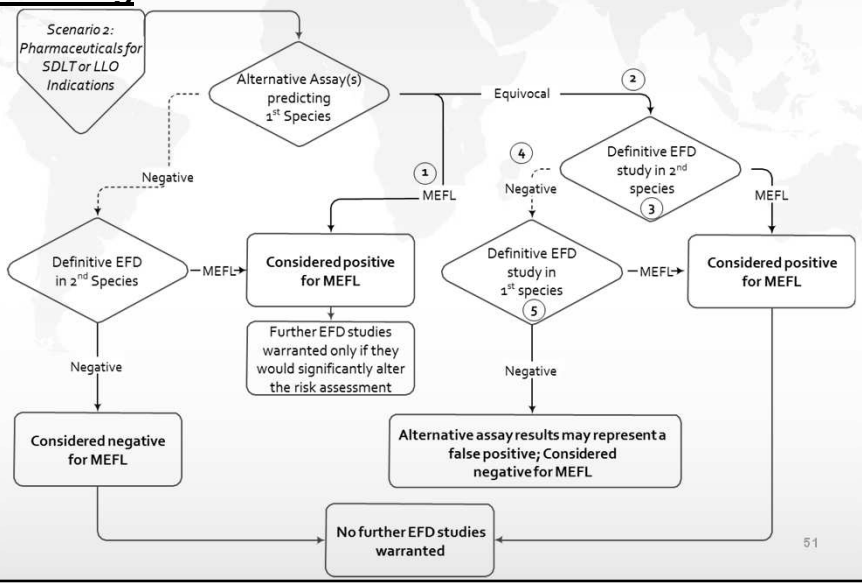



### *Pharmaceuticals <u>Expected to Be Embryo-fetal Toxicants</u>*

1. No additional assessment is warranted if unequivocal MEFL signal is observed at clinically relevant extrapolated exposures.
2. Alternatively, pEFD studies can be used; however, negative results should be confirmed by a definitive study in the relevant species
3. Conducting *in vivo* EFD studies in series, as shown, can permit reduction in animal use, as 2nd *in vivo* assay is not warranted if the first study is positive.

ICH S5(R3) – Step 4

## *Pharmaceuticals Intended to Treat Severely Debilitating or Life-Threatening Diseases*






ICH S5(R3) – Step 4

## Annex 2: Alternative Assays

### *Pharmaceuticals Intended to Treat Severely Debilitating or Life-Threatening Diseases – contd.*

1) A clearly positive MEFL signal at clinically relevant extrapolated exposures can be sufficient to consider a pharmaceutical positive for EFD toxicity, without further assessment, on a case-by-case basis.
2) While pEFD studies can be used, negative results from definitive *in vivo* EFD studies in two species are warranted to establish that alternative assay results represent a false positive.
3) For late-life onset diseases, given low likelihood of pregnancy in this patient population a pEFD study in the 2nd species can generally be sufficient.
4) Conducting *in vivo* EFD studies in series, as shown, can permit reduction in animal use, as 2nd *in vivo* assay is not to be conducted if the first is positive.
5) Same species as the alternative assay is intended to predict.

## Annex 2: Alternative Assays

### Examples of EFD Testing Strategies Utilizing AAs

***Pharmaceuticals Intended to Treat Late-life Onset Diseases***

- For diseases diagnosed in the older population, with low incidence in reproductively capable women
- Whether an EFD assessment is warranted needs to be determined on a case-by-case basis.
- Not intended for situations where the treatment population is presumptively infertile (e.g., post-menopausal osteoporosis), for which no EFD assessment would typically be warranted.
- Scenario is similar to that depicted for severely debilitating or life-threatening diseases, with the exception that the first *in vivo* assessment in the second species can be conducted as a pEFD study.



ICH S5(R3) – Step 4

## Annex 2: Alternative Assays

### Reference compound list

- Contains compounds that have been shown to induce MEFL in nonclinical studies and / or humans
- Only findings of MEFL were recognized for NOAEL and LOAEL determinations.
  - reversible or minor manifestations of developmental toxicity were not used for this assessment.
- The compounds in this list as well as others can be used to support qualification of an alternative assay or battery of assays.
- Negative controls are required to assess assay specificity.

## Summary of Guideline Content

**This revision R3 of ICH S5**

- Provides a focus on application to human risk assessment.
- Offers additional dose selection endpoints.
- Emphasizes the use of existing data.
- Describes integrated testing strategies for assessing DART, including for biologics.
- Outlines guidance on qualification of alternative assays for use in risk assessment for regulatory purposes.





## Considerations

- **In the interest of 3Rs the experimental strategy should minimize the use of animals, while not diminishing the overall human risk assessment.**
- **The revised guideline should be read in conjunction with ICH M3, ICH M3 Q&A, ICH S6 as well as ICH S9.**

## Guidelines for Implementation

- With the implementation of Annex 1 (*In vivo* study designs) and 2 (Alternative Assays) of the guideline, it is anticipated that industry will generate and submit more results from combined *in vivo* reproductive toxicity studies as well as outcomes from qualified alternative assays.
- This presents a driver for increasing regulatory experience on their applicability, and hence the potential for harmonization will increase.





## Conclusions

- For designing DART studies, all available pharmacological, toxicokinetic, and toxicological data for the pharmaceutical should be considered in determining which study design(s) should be used.
- ICH S5(R3) strengthens the role of 3Rs.
- ICH S5(R3) promotes novel testing paradigms for combined *in vivo* reproductive toxicity studies and alternative assays.

## Contact

- **For any questions please contact the ICH Secretariat:**

**admin@ich.org**